AD-A105 056 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL--ETC F/G 5/10
THE SOURCE SELECTION DECISION PROCESS IN AERONAUTICAL SYSTEMS D--ETC(U)
JUN 81 C W BARCLAY, J E NIDO
UNCLASSIFIED AFIT-LSSR 12-81 NL

1 of 2
AD
A105056

LEVEL I

3

AD A105056

DTIC ELECTE OCT 6 1981 S D A

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (ATC)

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

81 10 6 190

THE SOURCE SELECTION DECISION PROCESS
IN AERONAUTICAL SYSTEMS DIVISION.

Colin W. Barclay, Australian DOD
Jose E. Nido, Captain, USAF

AFIT-LSSR 12-81

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the Air Training Command, the United States Air Force, or the Department of Defense.

AFIT Control Number LSSR 12-81

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/LSH, Wright-Patterson AFB, Ohio 45433.

1. Did this research contribute to a current Air Force project?

   a. Yes   b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes   b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

   a. Man-years ________ $ ________ (Contract).

   b. Man-years ________ $ ________ (In-house).

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3 above), what is your estimate of its significance?

   a. Highly Significant   b. Significant   c. Slightly Significant   d. Of No Significance

5. Comments:

________________________   ________________________
Name and Grade   Position

________________________   ________________________
Organization   Location

FOLD DOWN ON OUTSIDE - SEAL WITH TAPE

AFIT/ LSH
WRIGHT-PATTERSON AFB OH 45433

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE. $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 73236 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFIT/ DAA
Wright-Patterson AFB OH 45433

FOLD IN

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER LSSR 12-81	2. GOVT ACCESSION NO. AD-A105 056
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) THE SOURCE SELECTION DECISION PROCESS IN AERONAUTICAL SYSTEMS DIVISION	5. TYPE OF REPORT & PERIOD COVERED Master's Thesis
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Colin W. Barclay, Australian DOD Jose E. Nido, Captain USAF	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Systems and Logistics Air Force Institute of Technology, WPAFB OH	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Department of Communication and Humanities AFIT/LSH, WPAFB OH 45433	12. REPORT DATE June 1981
	13. NUMBER OF PAGES 121
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Air Force Institute of Technology (ATC)
Wright-Patterson AFB, OH 45433

18. SUPPLEMENTARY NOTES

7 AUG 1981

APPROVED FOR PUBLIC RELEASE AFR 190-17

Fredric C. Lynch

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Source Selection
Proposal Evaluation
Procurement
Multi-Attribute Decision Making
Systems Acquisition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Thesis Chairman: Jack L. McChesney, Lieutenant Colonel, USAF

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This research is concerned with identifying a model of the source selection process as used in Aeronautical Systems Division, Air Force Systems Command (ASD) and evaluating the strengths and weaknesses of the process in relation to stated Department of Defense and Air Force objectives. Information was gathered from a review of past source selection cases and a series of interviews with ASD source selection personnel. A computer model was constructed to simulate the effects of the decision forming techniques observed on the possible outcomes of source selections. A resulting descriptive model provides a basis for better understanding of the quality of decision information provided by the process and forms a framework for improving the source selection process.

1473A

LSSR 12-81

THE SOURCE SELECTION DECISION PROCESS IN AERONAUTICAL SYSTEMS DIVISION

A Thesis

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Logistics Management

By

Colin W. Barclay, BTech, GradDipAdmin
Australian DOD

Jose E. Nido, BS
Captain, USAF

June 1981

This thesis, written by

Mr. Colin W. Barclay

and

Captain Jose E. Nido

has been accepted by the undersigned on behalf of the Faculty of the School of Systems and Logistics in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN LOGISTICS MANAGEMENT

DATE: 17 June 1981

Jack L McChesney
COMMITTEE CHAIRMAN

ACKNOWLEDGMENTS

We wish to thank Mr. James Schaeffer of ASD/PM for the benefit of his knowledge of source selection, and for his assistance in obtaining access to the source selection personnel and records necessary to be able to conduct this research. Thanks also are due to Mr. James Helmig and the staff of the ASD Source Selection Center for their cooperation and patience in making their records available to us. We are also deeply grateful to the many busy people in ASD source selection activities who gave generously of their time and accumulated knowledge.

Finally, our thesis advisor, Lt Col Jack McChesney, was an inspiration through his insight of the problems we faced, his helpful criticisms, and his encouragement. For this we thank him.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
I. SOURCE SELECTION	1
The Source Selection Decision Problem	1
Source Selection Decision-Making	1
Decision-Making Problems	3
The Research Need	4
Literature Review	4
Policy and Procedural Background	4
Purpose of Source Selection Procedures	10
Theoretical Background	11
Some Recent Propositions	16
Practical Considerations	17
II. RESEARCH APPROACH	20
Research Objectives	20
Scope of Research	20
Research Question	21
Research Methodology	21
Discussion	21
Research Hypothesis	26

CHAPTER | Page

III. ANALYSIS AND MODELING 29

Review and Analysis of Cases 29

Results of Regression Analysis 30

Modeling the Value-Building Process 32

Computer Model 33

Assumptions of the Model 34

Parameters in the Model 35

Goodness Level 35

Weighting Coefficients (B_i) 37

Introduction of B_o to the Model 38

The Computer Program 39

Analysis of Output of Computer Model . . . 41

Sample Size 43

ANOVA Test Procedure 44

ONEWAY ANOVA Results 45

Difference Between Color-Scored and Numerically-Scored Results (t-Test) . 49

Multiple ANOVA (MANOVA) 50

Summary of Results of Analyses of Model Output 52

IV. INTERVIEWS WITH SOURCE SELECTION PRACTITIONERS 55

Effectiveness and Efficiency of the Process 56

The Scoring Process 60

Contractor Inquiries and Deficiency Reports 63

Page

Improvements Suggested by Interviewees 65

Summary 69

V. CONCLUSIONS AND RECOMMENDATIONS 71

Source Selection Methodology 71

Maturity of Concept 75

Weights of Attributes 76

Source Selection Resources 77

Management Style of the SSA 77

Choice of Scoring Method 79

Procedural Aspects of Source Selection . . 81

Effectiveness and Efficiency 82

Use of Scoring Techniques 83

Contractor Inquiries 84

Problems of Source Selection 85

Recommendations 86

APPENDICES 89

A. LISTING OF COMPUTER PROGRAM 90

B. SUMMARY TABLES OF COMPUTER OUTPUT 94

C. OUTPUTS OF MANOVA TESTS 105

D. GUIDE TO INTERVIEWS WITH SOURCE SELECTION PRACTITIONERS 114

SELECTED BIBLIOGRAPHY 117

A. REFERENCES CITED 118

B. RELATED SOURCES 121

LIST OF TABLES

TABLE Page

I. GOODNESS LEVELS OF SIMULATED PROPOSALS 36

II. POSSIBLE SETS OF VALUES OF WEIGHTING COEFFICIENTS 38

III. RESULTS OF ONEWAY ANOVA TESTS 47

IV. TABLE SHOWING HOMOGENEOUS SUBSETS OF TREATMENTS 48

V. T-TEST OF RESULTS USING NUMBER AND COLOR SCORE METHODS 51

LIST OF FIGURES

Figure Page

1. Source Selection Value Hierarchy 23

CHAPTER I

SOURCE SELECTION

The Source Selection Decision Problem

Air Force system acquisition projects involve contracts at stages throughout the project life for the procurement of services and equipment toward fulfilling the system acquisition objectives. Typically, each procurement involves the solicitation of offers followed by an evaluation process which produces information from which a choice is made to determine the contract award.

Source Selection Decision-Making

The evaluation leading to the award decision (source selection) is the composite product of the results of independent expert assessments of a variety of aspects of the offers under consideration. While these aspects may cover a wide range of criteria, they can be conveniently grouped into five major areas: (1) technical, (2) operational, (3) logistics, (4) management, and (5) cost. Each area may in turn be broken out into more specific items which themselves may shred out into more discrete segments called "factors" (17:pp.3-2,3-3). The regulatory documents which are reviewed in this study give broad guidance on

the use of subjective and objective methods of integrating expert assessments of the separate aspects of offers into consolidated evaluation reports to provide a basis for the source selection decision. The evaluation should provide a balanced appraisal of all significant factors with a high level of quality and consistency to facilitate an objective, impartial, equitable, and economic comparative analysis of competing offers.

Determining an integrated evaluation of an offer from factor assessments involves two processes: (1) scoring and (2) weighting.

(1) Scoring is the allocation of a comparative "value" to each factor being assessed. The value may be expressed by allocating a numerical score, by color coding, by ranking, by narrative, or by a combination of these methods.

(2) Weighting is the process of giving to the score value of each factor a coefficient which reflects the relative importance of the factor in the final evaluation. In practice, weighting may be done by objectively allocating numerical coefficients or by subjective comparisons.

As the *foregoing suggests*, a number of empirical models have been developed as a guide in making evaluations for source selection decision-making. Because of the necessarily large number of participants in factor assessment and the variety of integrating techniques available, it is difficult to justify that current procedures provide the required objectives of quality and consistency in Air Force source selections.

Decision-Making Problems

A recent study (10:49) of source selection procedures in Aeronautical Systems Division (ASD) showed that different groups of source selection evaluators rated the same proposals differently. An examination (8:119) of numerical scoring and weighting schemes concluded that small relative changes in item weights and item scores can overturn the order of evaluations when differences between scores are small fractions of the scores. Awareness of shortcomings in numerical rating schemes has led to a "strong trend toward rating proposal elements using a combination narrative and color coding system [1:pp.9,10]." While this trend avoids the specific criticisms of numerical rating schemes, there is no hard evidence to show that more subjective methods than numerical weighting come closer to obtaining consistent and accurate source selection decisions.

The Research Need

There remains a need for a source selection evaluation procedure which is capable of giving demonstrably consistent results with different assessment groups. Source selection from among complex competing offers is a multi-attribute decision-making situation.

Some recent academic discussions of possible applications of multi-attribute decision theory to military logistics problems are discussed in the literature review. It is believed that theory developments in this field offer a base from which to examine an actual source selection process to determine a source selection decision-maker's visibility of the assessment factors in relation to the source selection criteria. A sufficiently rigorous examination may provide new insights into source selection that will enable the development of more effective management of the process.

Literature Review

Policy and Procedural Background

Department of Defense Directive (DODD) 4105.62 (20:2) provides source selection policy and procedures for the acquisition of major defense systems, and states three primary objectives to be met as a result of the source selection process.

> The prime objectives of the process are to (a) select the source whose proposal has the highest degree of realism and credibility and whose performance is expected to best meet Government objectives at an affordable cost; (b) assure impartial, equitable, and comprehensive evaluation of competitors' proposals and related capabilities; and (c) maximize efficiency and minimize complexity of solicitation, evaluation and the selection decision.

The major systems acquisition process is a complex one. It consists of a sequence of specified phases of program activity and decision events directed to the achievement of program objectives in the acquisition of defense systems. Each major weapon system acquisition program has its unique features, and therefore, no two programs are identical. If one were to compare various programs, a number of differences would immediately surface to include differences in time, cost, technology, management, and contracting approach. Despite the differences, however, the basic acquisition process is common to all programs. As such, all programs are driven through the process toward a common goal of obtaining for the Government "the most advantageous contract--prices, quality and other factors considered [18:p.1-302.2]."

DODD 4105.62 (20:2) provides guidance to achieve this objective:

> Each DOD Component shall develop, and consistently apply, procedures which create the environment for an impartial, balanced and realistic appraisal of all proposals submitted.

Air Force Regulation (AFR) 70-15, <u>Source Selection Policy and Procedures</u>, (17:p.1-1) establishes policy, assigns authority and responsibilities, and prescribes implementing procedures for source selection. It also states the main objective of the source selection process:

> The prime objective of proposal evaluation and source selection is to assure impartial, equitable, and comprehensive evaluation of competitive proposals and to assure selection of that source whose proposal, as submitted, offers optimum satisfaction of the Government's objectives including cost, schedule, and performance.

A typical source selection process is composed of a structured organization which consists of a Source Selection Authority (SSA), Source Selection Advisory Council (SSAC), and a Source Selection Evaluation Board (SSEB) (17:p.1-4). The source selection process itself is initiated as a result of the submission and approval of a Source Selection Plan. This plan, the key planning document for the conduct of the source selection process, is normally prepared by the project officer charged with effecting the procurement of the system (17:p.2-1). The Source Selection Plan usually includes, among other things, "basic evaluation criteria to provide a basis for the more detailed shredout by the SSAC and SSEB for use in the solicitation," a description of the SSEB evaluation and rating methodology and the SSAC analysis technique, and a schedule of events, identifying and listing the source

selection activities within a time framework (17:pp.2-1, 2-2).

The Contract Definitization Group (CDG) is a part of the SSEB organization. Its role is to negotiate definitive contracts with all offerors determined to be in the competitive range. The CDG manages all communications with the offerors and is advised by a Cost Panel. The primary purpose of the Cost Panel is to provide an evaluation of the most probable cost to the Government of each offeror's proposal (1:16). While technical and cost evaluations by different evaluators are held simultaneously, they are kept apart to prevent the technical evaluation from being biased by cost considerations.

After proposals are received, the evaluation period commences with the SSEB examining and conducting:

> . . . an in-depth review of the relative merits of each proposal against the requirements in the solicitation document and the evaluation criteria established by the SSAC. The evaluation function must be thoroughly conducted, objective, fair, and economical [17:p.1-6].

A summary report of findings by the SSEB is then prepared and submitted to the SSAC. This SSEB Evaluation Report is basically a summary of the results obtained after evaluating each proposal against the standard criteria set forth by the SSAC (17:p.1-6).

AFR 70-15 (17:pp.1-5,1-6) establishes the SSAC's duties and responsibilities. These include, among others:

> (1) Establish the evaluation criteria, using the general guidance set forth in the approved Source Selection Plan.
>
> (2) Establish the relative importance of the evaluation criteria in a form for use in the solicitation document.
>
> (3) Establish the evaluation criteria weights for SSAC use when numerical scoring techniques are employed.
>
> (4) Review the findings of the SSEB and, when numerical scoring has been used, apply the established weights to the evaluation results.
>
> (5) Prepare the SSAC Analysis Report (comparative analysis) based on the SSEB Evaluation Report.

Basically, this part of the process consists of a review of the SSEB Evaluation Report by the SSAC, after which an evaluation of proposals is again conducted against the SSAC criteria. A Source Selection Advisory Council Analysis Report is then submitted to the SSA. This comparative analysis report consists of a "proposal versus proposal" evaluation that should help the SSA make an objective selection decision (17:pp.1-1,1-2,1-5,1-6).

The SSA is ultimately responsible for the proper conduct of the proposal evaluation and source selection process. Therefore, he should strive for a source selection process that will provide him with the information necessary to make the most objective selection decision

possible.

> The SSA must be presented sufficient indepth information on each of the competing offerors and their proposals to make an objective selection decision. The SSAC Analysis Report and oral briefing should be presented to the SSA in a manner which accomplishes this objective. The SSAC presents findings and analyses but does not make recommendations to the SSA unless specifically requested [17:p.1-2].

In the final analysis, the degree of success that the SSA will attain in making an objective decision will depend on the extent to which a logical, consistent, and systematic approach is established. AFR 70-15 (17:p.1-3) provides guidance for the establishment of evaluation criteria and rating systems to be used in evaluating offerors' proposals:

> The specific evaluation criteria must be included in the solicitation document and enumerated in terms of relative order of importance of those significant factors which will form the general basis for proposal evaluation and selection/contract award . . . The rating system shall be structured to enable the SSA to identify the significant differences, strengths, weaknesses, and risks associated with each proposal and subsequent definitized contract . . . The rating system may be entirely narrative, or may employ numerical scoring and weights or a descriptive color code in conjunction with narrative assessments. The important task in either rating system is the integrated assessment of all aspects of the evaluation, analysis, and negotiation process.

AFR 70-15 (17:p.3-4) relies on the evaluator's own judgment while performing an evaluation:

> How an evaluator approaches the task of evaluation is up to his own judgment based on his experience. The method by which it is accomplished is dependent on what he feels best suits the particular circumstances . . . It is, however, important that all evaluators be consistent in their approach to evaluation. Failure to do so will result in distortion of the true value of the proposals.

Purpose of Source Selection Procedures

The Logistics Management Institute briefed the Defense Blue Ribbon Panel on the subject of defense procurement policy and weapon systems acquisition in August 1969:

> Formal procedures were established for selecting contractors for major development or large production efforts. These procedures required evaluation of proposals according to pre-established, point grading criteria and a review of the documented results of the grading system. The objective was to reduce the influence of subjective judgments in the selection of contractors and to encourage objective evaluation of all proposals by responsible offerors [9:21].

The essential decision-making process in source selection involves weighing and judging complex issues arising from the assessment of the many factors which make up competing offers. The issues are evaluated by separate expert groups with different perceptions of the ultimate acquisition. Weighting of issues is subject to the biases of the weighters. Overall policies may be overwhelmed by the goals of the organizational subsystems involved in the process. The source selection decision-maker requires information which:

(1) relates to the acquisition policy and objectives

(2) is free from bias

(3) is equitably weighted

(4) can withstand scrutiny and be repeatable with different assessors.

Finally, the decision-maker requires the information in a form which is digestible and which will assist him to exercise judgment in the fullest possible knowledge of the choices available.

Theoretical Background

A preliminary survey of general literature in decision-making suggests that there are research findings which might be applied to the experience of existing empirical source selection models to develop an improved understanding of the source selection process.

Simon (16:272) proposed the concept of bounded rationality as a feature of management decision-making. He reasoned that decision-makers in complex situations "satisficed" the choices available to them. They used only that part of the available information which they perceived to enable them to make a satisfactory rather than an optimal decision. The cautions and reservations expressed in current Air Force and DOD source selection regulations

confirm an awareness of the difficulties of source selection decision-making. Because of this inherent complexity, "satisficing" continues to play a significant part in source selection decision-making as a practical necessity.

These views have support in a recent research which utilized multiple linear regression techniques to examine source selection in an Air Force procurement division. Milligan (10:vi) attempted to determine whether or not the evaluation criteria contributed significantly to the rating a proposal received and how program managers and supervisors make source selection decisions. He found that people do not always use all the information available to them in making source selection decisions:

> Thesis results suggest that source selection decisions are not similar across organizations within the AF division. Furthermore, subjects did not utilize all the information available to them in making decisions. People often chose to utilize only a part of the available information in arriving at a decision.

Dawes (3:180-188) demonstrated the "bootstrap" effect of making policy a conscious element of the decision model. When policy was defined or "captured" in the model, decisions became more aligned to policy. A similar effect in Air Force source selection decision-making was suggested by Milligan when he showed that source selections were more consistent among experienced source selection staff when given a policy in the decision task statement than when

they were given the task with no formal policy.

While trying to improve the proposal evaluation phase of the source selection process, Dycus (5:256) conducted an Evaluator Preference Survey in which he attempted to measure the attitudes and evaluative judgment of a quasi-sample of 33 experienced DOD technical proposal evaluators. Although he found that the evaluators had a favourable attitude toward proposal evaluations, survey responses indicated a need for improvement of the evaluative procedures:

> . . . survey data indicated considerable room for the government to improve proposal evaluation mechanics. Most evaluators indicated they reinterpreted scored evaluation criteria. There was only moderate judgment that scores evaluation criteria and rating scales were "good" and "fair".

Dycus recommended that experimental research be conducted in order to improve the proposal evaluation aspect of source selection. He further suggested that such research would improve evaluation rating scales, evaluation criteria for scoring, determine preferred evaluation mechanics, and improve scoring discrimination:

> End product of proposed experimental research would be a proposal evaluation guide that defines a preferred rating scale, and directs the evaluators in how to make their evaluation scorings. Such a guide would improve the quality and discrimination of proposal evaluation scores, and attest to the practical value of applied procurement research [5:256].

The primary goal of source selection is to arrive at an objective selection decision. However, several problems exist which limit the ability to accomplish this goal. The work of Beard (2:iv) in his study, "The Application of Multi-Attribute Utility Measurement (MAUM) to the Weapon Systems Source Selection Process", identifies five problem areas that presently limit the ability to fully accomplish an objective evaluation:

> These problems are: current weapon systems development is multidimensional and does not allow for evaluation on a single dimension - an array of attributes must be evaluated; performance evaluation is in many cases a subjective attribute and judgment can be influenced by biased viewpoints; the current color coded evaluation procedure provides results that can be washed out and are arrived at wholistic; the current numerical evaluation procedure provides results that can be very close, may tend to level out results or obscure the more important issues; and costs.

MAUM is a ten-step procedural approach developed from multi-attribute utility theory by Dr. Ward Edwards to objectively address important decisions when selecting among various alternatives having multiple attributes (technical, logistic, and operations evaluation factors) (2:8). It provides a framework for scoring and weighting attributes in such a way as to ensure significant discrimination between the scores allocated to substantially different proposals. Using this approach, Beard (2:43) concluded that the objectivity required in source selection decisions can be attained:

> MAUM's proceduralized methodology greatly reduces the influence individual bias can have in evaluation results. The use of value curves and the philosophy of "operationally defining" evaluation factors will result in much more objective evaluations. MAUM's procedures preclude inconsistent application of evaluation standards over time.

Basically, Beard argued that since the present source selection evaluation process considers various proposals having multiple attributes (evaluation items and factors), MAUM's ability to evaluate decisions having more than one attribute, aspect, criterion, or dimension helps eliminate the problems presently encountered in the process (2:42).

There has been concern regarding the numerical scoring and weighting system used to evaluate offerors' proposals, specifically, the sensitivity of total scores to small variations in the choices of item weights and in item scores. In a paper presented to the Sixth Annual Procurement Research Symposium, Lee was concerned with the possibility of these variations causing "offeror A to have a greater total score than offeror B in one case, while making B's score exceed A's in another (8:123)." Lee concluded that:

> The order of numerical scores of proposals can be overturned by small relative changes in item weights and item scores whenever differences between scores are small fractions of the scores, even when item weights meet all the requirements of AFR 70-15 and AFLC Supplement 1 to that regulation.

Some Recent Propositions

Waid (21:12) has described DOD material acquisition decision-making as a value-building process and proposed the use of a theory of analytic hierarchies to clarify the multi-criteria choice situation involved. A key element of the theory is the use of two-dimensional comparison matrices to refine expert estimates of value scores in terms of the value structure of the organization. A claimed advantage of the scheme was that it is relatively simple to construct and administer in a complex organization. He concluded that the process drives toward consensus and provides a truly wholistic approach to decisions.

A proposed application of multi-criteria decision theory to a specific Air Force acquisition planning decision-making scenario by DeWispelare, Sage, and White (4:p.1-15) identified two major theoretical application techniques: Multiple Objective Optimization Techniques (MOOT) and Multiple Attribute Utility Theory (MAUT). It was suggested that both MOOT and MAUT are mental constructs to approaching multiple criteria decision situations and that there are practically no fundamental differences between their analytical structures. However, while MOOT may more quickly identify non-dominant solution sets, decision-maker preference (weighting) emerges more

efficiently through MAUT. Dewispelare, Sage, and White developed a methodology of combining the organizationally desirable features of MOOT and MAUT. The methodology has been tested in the Air Force and has been demonstrated to be an acceptable and desirable approach to improving the efficiency of decision-making. The research offers encouragement of the practicality of developing the application of multi-criteria decision theory to source selection scoring and weighting.

Practical Considerations

The literature review to this stage suggests that the major problems in achieving effective source selection evaluations include consistency, equitable weighting of factors, and policy visibility. Recent research has focussed on methodology for improving the quality of estimates of value score and attribute weighting.

In general, source selection practitioners are averse to the use of mathematical models of subjective judgment which use numerical scoring (for example, (6:89)). There is a feeling that numerical scoring methods inhibit the freedom of the decision-maker to make, and justify, subjective decisions. During recent years, there has been a strong trend toward using a combination narrative and color-coding system in rating proposal elements. When

used, the following is an example of how the color-coding system may be applied (1:10):

- Blue - Exceeds specified performance or capability and excess is useful, high probability of success, no significant weaknesses.
- Green - Average, meets most objectives, good probability of success, deficiencies can be corrected.
- Yellow - Weak, low probability of success, significant deficiencies, but correctable.
- Red - Key element fails to meet intent of Request for Proposal (RFP).

The source selection evaluation process provides for a tendency to "wash" the evaluation of proposals toward an acceptable standard. This effect appears to be due in part to the conservativeness of evaluators at the lower levels. These evaluators appear to be reluctant to rate a proposal as unacceptable and so eliminate it from the competition. Evaluators at these levels seem to avoid this kind of decision, deferring it to the higher level to make such a determination. The cumulative effect reduces the visibility at the higher levels of the process of the overall worth of the different proposals when compared against standards.

It is believed that detailed study of the value building processes in actual source selections is necessary before significant conclusions can be made about the

practical application of multi-attribute decision-making theory to improving source selection.

CHAPTER II

RESEARCH APPROACH

Research Objectives

The objective of this research was to examine the value building processes in an actual source selection case and to establish its correspondence with the theoretical constructs of multi-attribute decision theory.

Scope of Research

Formal source selection evaluation procedures are mandatory only for new development programs requiring $100 million or more RDT&E funds or projected to require more than $500 million production funds, or other programs specifically designated (19:2). However, the objectives of source selection remain the same in all procurements regardless of dollar value, and responsible officers are required to demonstrate a systematic and consistent approach to the source selection decision. The problems of offer evaluation and decision-making are similar for all procurements and vary only in size and scope.

This research was directed to a detailed study of a source selection process undertaken within the standardized formal procedural framework used within the

Aeronautical Systems Division (ASD). However, it is considered that it will provide a basis for study of more general cases of government source selection.

The research included a series of interviews with source selection practitioners and administrators to identify perceptions of the strengths and weaknesses of the empirical source selection models in current use.

Research Question

The question addressed in the study introduced by this paper may be summarized in the following way:

> Can a detailed study of an actual source selection process establish a relationship with theoretical multi-attribute decision-making models which will provide means for improving the management of source selection?

Research Methodology

Discussion

The process of source selection may be pictured conveniently as a value hierarchy in the manner described by Waid (21:14). In source selection, the levels of hierarchy are typically:

source
area
item
factor
sub-factor

The values of each component which contribute to the decision rise progressively through the hierarchy, successively being refined, until they reach the source decision level. At each level, a series of weighted combinations of individual component scores takes place to arrive at a new set of scores to enter the weighting process at the next higher level (Figure 1).

Researchers have modeled this kind of combinational process in many applications as a linear multiattribute utility model (22:122-124). The model is expressed in the form:

$$Y = B_1x_1 + B_2x_2 + B_3x_3 + \cdots + B_nx_n$$

where Y is the value (score) outcome of the process, and x_n is the value of the nth component, and B_n is the weighting coefficient of the nth value.

The linear model is consistent with the procedures outlined in AFR 70-15 (17:p.2-6) and the source selection policy-capturing research of Milligan (10:12).

The underlying assumptions of the linear model are that the components are (22:123):

(1) independent - to avoid double counting, and

(2) unidimensional - the scores should be realistically seen as adding to the decision dimension, and

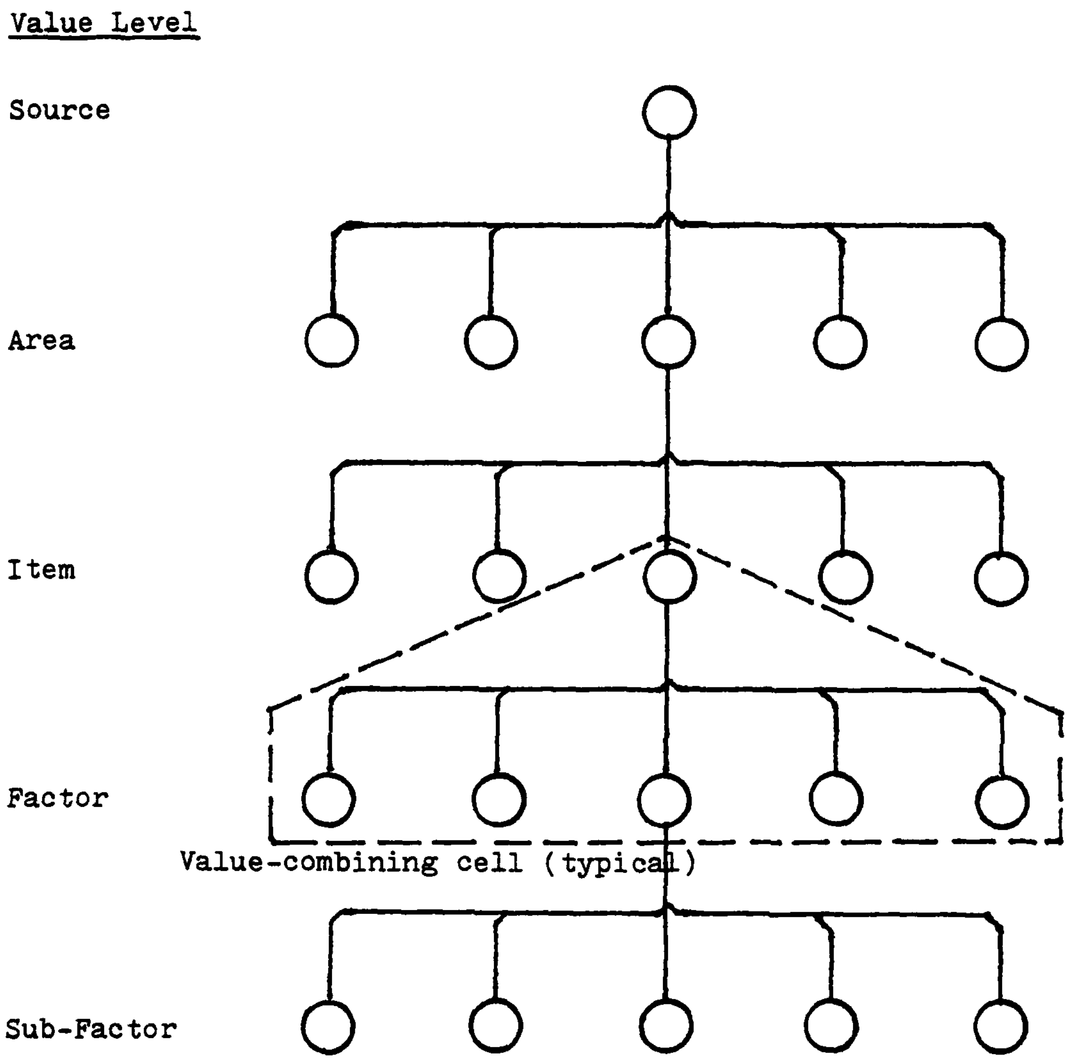


FIGURE 1 - Source Selection Value Hierarchy

(3) compensating - high scores on some components will compensate for low scores on others, and

(4) relevant - the components should be relevant over all contexts, and

(5) exhaustive - all appropriate components should be included, but

(6) determinant - the components should be important to the selection.

When comparing a number of competing multi-attribute options (proposals) as in source selection, meaningful comparisons are made when the coefficients B remain constant for each calculation of Y.

This kind of model implies a straightforward way of combining objective component scores through a value hierarchy to evaluate competing proposals. The combined scores become absolute comparative values at the source level which should provide a clear basis for the source selection decision.

However, numerical scoring systems have been almost universally rejected in ASD as a suitable means of source selection for anything but the simplest procurements.

The major objections to numerical scoring arising out of practical experience are:

(1) variations between offers are "averaged out" so that major deviations from standard become obscured in the final score.

(2) The allocation of objective scores narrows the options of the decision-maker constraining him to the highest numerical score.

These objections are supported in a stated reluctance of evaluation personnel to use the full range of numerical scoring scales and the findings by Lee (8:123) that when total scores are close, small variations in component scores can overturn the result. The latter gives concern that the former can lead to a scoring result that is not optimal.

ASD prefers the use of more subjective color coding scoring systems which are believed to give the combiner at each level in the value hierarchy greater flexibility in the subconscious weights that he applies to subordinate component color scores when allocating his own color score to the whole of the group of components within his responsibility.

In most important acquisitions, the potential contractors are experienced and competent in government contracting. They have developed an intimate knowledge of the government's requirements during prior negotiations and understand the source selection process. Their

proposals are therefore constructed to closely conform to the expected criteria. The outcome is that all proposals tend to meet the main acquisition requirements and that differences between them are small. Differences between proposals tend not to be evident in the comparative color score allocated at higher levels in the value hierarchy. The Source Selection Advisory Council is often obliged to look below the highest hierarchical levels to detect differences which may become justifiable bases on which to make accept/reject decisions. It appears in practice that at each level of combination of scores there is an effect which "washes-out" the visibility of significant factors which may be a basis for acceptance or rejection when considered with the whole.

Research Hypothesis

Even with the use by ASD of color-coded scores in the value hierarchy, the concept of a linear combinatorial model remains valid if meaningful numerical equivalent scores may be given to the color code. However, because the color-coding technique of scoring does not lend itself to a priori allocation of objective attribute weights (other than a simple ranking of order of importance), it is possible to bias the model during application. The bias effect may be represented simply by extending the model by

a constant term B_o, such that:

$$Y = B_o + B_1x_1 + B_2x_2 + . . . + B_nx_n$$

AFR 70-15 suggests a suitable scale of numbers in the range 0-10 to equate to color scorings. This scale was adopted as a basis to score color-code ratings used in source selection:

blue	-	10
green	-	6
yellow	-	2.5
red	-	0

The value-building processes in actual source selection cases were examined to see if a fit could be established between the actual processes and the extended model. A survey of recent ASD source selections showed that suitable cases could be identified with sufficient proposals and components to be able to conduct a multiple regression analysis of the allocated cell value to the component values for each cell of the decision hierarchy.

Analysis was conducted by the multiple regression procedures in the Statistical Package for the Social Sciences (SPSS), (12:328), available on the Cyber CDC 6600 Computer. The basic test hypothesis was:

$$H_o : B_o = 0$$

$$H_1 : B_o \neq 0$$

Having established the possible nature of the value-building equation when color-scoring was used as the

discriminant, a computer model was constructed to simulate a series of value-building situations to examine the relative performance of color-scoring and numerical scoring methods as value discriminants.

Interviews were held with experienced source selection practitioners in ASD to gain a fuller appreciation of the empirical source selection process and to validate the model assumptions.

The outcome enabled some conclusions to be made about the way the source selection process functions and its relationship to the DOD objectives.

CHAPTER III

ANALYSIS AND MODELING

Review and Analysis of Cases

A preliminary survey of source selection cases completed in ASD over the period 1975 to 1980 was made to identify cases with sufficient historical data of the value building process to facilitate detailed analysis of value building hierarchies.

Three cases were identified as being suitable for analysis, and permission was granted by the Commander, ASD to examine the records in detail. The selected cases involved teams of 41, 43 and 70 evaluators and 4, 5 and 9 proposals respectively. All cases used combinations of color and narrative scoring techniques.

Within each case, it was sought to isolate individual value building cells which met criteria for multiple regression analysis.

The criteria sought for analysis were:

(1) The higher-level composite "value" given to a value building cell was expressed in comparative terms to the expression of values of the component parts or attributes (i.e.,

colors, or by generic descriptive groupings such as "Exceeds Standards", "Meets Standard", "Fails to Meet Standard", "Unacceptable").

(2) There were a sufficient number of proposals in relation to the number of independent components so that the multiple regression synthesis of relationship was valid, i.e., the number of attributes (variables) was less than the number of proposals (sample size) (12:329).

Twelve value building cells were identified which met the criteria.

Results of Regression Analysis

The twelve value cells examined had from three to thirteen components. However, where all proposals scored the same for a component, that component was eliminated from inclusion in the equation as being discriminating. The SPSS multiple regression technique was then applied to evaluate the relationship of implicit value:

$$Y = B_o + B_1x_1 + B_2x_2 + . . . + B_nx_n$$

Further components were eliminated in the regression analysis because of multi-collinearity or because they were below a 0.01 inclusion level. As a result, all value cells reduced to five or less significant components in

the analysis. B_o was found to be equal to zero with 95% confidence in only two of the twelve relationships so derived. Values of B_o identified with 95% confidence were two in the range 0 to 7.7 and eight in the range -10.3 to 0.

The B_i values reflect the evaluator's perception of the relative importance of the components of the decision cell. The structure of the source selection process requires that weightings be determined in advance of and separately from the component evaluations (17:p.3-7). It is difficult, if not impossible, for the evaluator to express absolute weights in numerical terms when using color or narrative scoring. The observed practice is only to rank components in order of relative importance to each other at the outset and it is the judgment of the evaluator which determines the implicit relative weight actually accorded to each component at the time of final determination of the composite score.

The value of B_o may be perceived as a measure of the cell evaluator's adjustment of the weighted composite score against a subjective benchmark. It reflects a subjective readjustment of the value of the competing proposals in an attempt to portray a relationship between them and the perceived standard.

The composite scorer for the cell, therefore, undergoes a complex process of mental weighting and re-

evaluation of the component score data in arriving at a value. When using the color code/narrative approach of ASD, he is constrained to express the judgmental outcome by one of four discrete "values" (red, yellow, green or blue), shaded as necessary by narrative support.

Modeling the Value-Building Process

Dr. Lee has shown (8:119) that when numerical scoring schemes are used, the order of numerical scores of the whole are sensitive to small relative changes in component weights and scores whenever differences in scores are small.

This part of the research was concerned with how the sensitivity of the model was affected when a four-increment scale of scoring was used instead of a relatively continuous numerical scoring scale; and to see what effect the introduction of the B_o value adjustment had on the discriminating power of the model.

In considering the discrimination between different proposals when color-scoring is used, it was evident that a difference becomes significant when component scores are near the "border-line" of an incremental range on the scoring scale. Because of the discontinuous nature of the discriminating effect of score differences, it was decided that the problem could be most conveniently

examined by means of a computer simulation. A computer model was therefore constructed to find out the effectiveness of the scoring system used by ASD as a means of discriminating between offers of various differences, and to compare the performance of color scoring with numerical scoring.

Computer Model

The model was constructed to simulate a value building cell of five components, the whole value of which is represented by Y where:

$$Y = B_o + B_1x_1 + B_2x_2 + B_3x_3 + B_4x_4 + B_5x_5$$

The values, x_i, of the components of the cell were randomly generated for five value cells, representative of the situation of evaluating five competing proposals. The data were generated to represent five sets of proposals of differing degrees of "goodness" so that the discriminating properties of the model might be observed. For each set of five lots of simulated data, the five item values (Y) were calculated and the highest scored proposal was determined. Multiple sets of data were tested over a range of values of B_o and B_i and goodness levels to determine the frequency of selection of the "best" proposal for each set of independent variables. The model also replicated the process using the raw numerical scores of x_i

instead of the four-increment color scoring scale.

Assumptions of the Model

Examination and analysis of the case histories suggested that the following were reasonable assumptions on which to base a model synthesis:

(1) Each component item of the value cell is independent, i.e., no multi-collinearity exists.

(2) The value attributed to each proposal in the whole may be conceptualized on a scale of 1 to 10 and that the limit of perception of objective difference of values of proposals so conceptualized is 2 per cent. If the objective difference is less than 2 per cent, then the "best" bid will be selected on subjective factors.

(3) Goodness has consistency. A proposal for which the evaluation is "good" in an item may be expected to perform at a "good" level on the average across all factors that make up the components of the item evaluation.

(4) Evaluators tend to judge components against the standard on a continuum before allocating discrete color or descriptive scores.

Parameters in the Model

Goodness Level

In order to be able to examine the discrimination of the model, it was necessary to simulate data representing the evaluated component scores of proposals of differing quality or "goodness".

Assumption 3 states that the values attributed to the components of a "good" proposal in a value cell will cluster about a value higher than the value about which bids of lesser goodness will cluster. To simulate this concept, "goodness" levels were modeled on a scale of 1 to 10. The designated "goodness" level was set as the mode of a continuous triangular frequency distribution. A computer-generated, uniformly distributed pseudo-random number was then put against the cumulative distribution curve of the triangular distribution to derive a "goodness" number. The numbers (AX) so derived were then reduced to a scale of color-equivalent incremental values (X) as listed in AFR 70-15 (17:p.3-6) as follows:

```
If (AX.LT.1.25) then X=0
If (AX.GE.1.25.and.AX.LT.4.25) then X=2.5
If (AX.GE.4.25.and.AX.LT.8) then X=6
If (AX.GE.8) then X=10
```

Sets of five proposals were simulated with each proposal in the set being of a designated "goodness" level. Each proposal consisted of a value cell with five components.

In each set of five proposals the "best" proposal was put at a goodness level of 10 and goodness levels of the remaining proposals put at 10 per cent decrements. In each successive goodness set, the interval between the "best" and the "second best" proposal was increased by 10 per cent.

The resulting goodness levels of the proposals in the sets are shown in Table I.

TABLE I

GOODNESS LEVELS OF SIMULATED PROPOSALS

Proposal Set	Proposal Number 1	2	3	4	5
1	10	10	9	8	7
2	10	9	8	7	6
3	10	8	7	6	5
4	10	7	6	5	4
5	10	6	5	4	3

Proposal sets were not extended beyond set number 5 because:

(1) it was judged that a 10:6 quality ratio was representative of the largest gap between proposals which would merit formal source selection procedures, and

(2) the difference between the "best" and "second best" proposal could no longer be regarded as "small".

Weighting Coefficients (B_i)

When total value is determined by the expression

$$Y = B_1x_1 + B_2x_2 + B_3x_3 + \ldots + B_nx_n \quad \ldots \quad (1)$$

and Y and x_i are both scored on the same value scale, then:

$$\sum_{i=1}^{n} B_i = 1 \quad \ldots\ldots\ldots\ldots \quad (2)$$

Typically (13:33), weighting coefficients when used in source selection are put at values which are multiples of 0.1. Within these guidelines, there are seven possible sets of values for B_i when five terms are included in the total value expression (Table II).

TABLE II

POSSIBLE SETS OF VALUES OF WEIGHTING COEFFICIENTS

Set Number	Value of B_i
1	0.6, 0.1, 0.1, 0.1, 0.1
2	0.5, 0.2, 0.1, 0.1, 0.1
3	0.4, 0.3, 0.1, 0.1, 0.1
4	0.4, 0.2, 0.2, 0.1, 0.1
5	0.3, 0.3, 0.2, 0.1, 0.1
6	0.3, 0.2, 0.2, 0.2, 0.1
7	0.2, 0.2, 0.2, 0.2, 0.2

All seven possible sets of B_i values were included in the computer model.

Introduction of B_o to the Model

The analysis of twelve value-building cells from three source selection cases revealed B_o values ranging from -10.3 to +7.7. The number of cells examined is a small sample compared with many source selection cases. Given the small sample size, it was not possible to make significant conclusions about the real limits of range and frequency of occurrence of B_o values when color-scoring or narrative-scoring systems of value expression were used. However, it was sufficient for this research to observe the possibility of occurrence of significant

B_o values. When a real value of B_o was introduced into the value equation and all B_i values sum to 1, as expressed in equation (2), it was necessary to modify equation (1) to retain the same scoring scales for Y and x_i, so that:

$$Y = B_o + (1-B_o/S)(B_1x_1+B_2x_2+B_3x_3 + \dots +B_nx_n) \dots (3)$$

where S is the scoring scale for Y and x_i.

Since the concept of the model was that the values of the components (x_i) were additive toward the value of the whole, and as B_o approached S the value of Y approached B_o, the maximum practical limit of B_o was S.

For the purposes of the model, three values were chosen for the adjustment parameter:

$$B_o = +7$$

$$B_o = 0$$

$$B_o = -7$$

as a basis to observe the effects of inclusion of B_o in the value building equation.

The Computer Program

The computer program to simulate the operation of the model is listed at Appendices A1-A3.

The program was arranged to give 80 simulations of data for each goodness set, providing 2000 simulated data points. The data were processed to find the value or score for each bid/goodness set combination and to select

the highest scoring bid for each simulation. The proportion of each bid selected over the 80 simulations was calculated. The outputs which the program provided were:

(1) A frequency table for each goodness set of per cent each proposal selected in the first run of simulations for the five goodness sets and three values of B_o and 7 sets of B_i coefficients.

(2) A histogram for each frequency table.

(3) A summary table of the frequency of selection of the "best" proposal (bid number 1) for each run of simulations against goodness set, B_o value and B_i coefficient set.

The program was arranged to run the 80 simulations five times, each time from a new random number base. The five runs were repeated using absolute numerical scores for discrimination instead of incremental color scoring to provide a basis for comparison between the two scoring methods.

The ten summary tables are presented at Appendices B1-B10 which show, for the same proposal data:

(1) Frequency of selection of proposal number 1 against goodness and B_o and B_i sets for color scoring for five runs.

(2) Frequency of selection of proposal number 1 against goodness and B_o and B_i sets for numerical scoring.

Analysis of Output of Computer Model

The computer model experiment was intended to study the effects on the numbers of times the "best" proposal was selected by successively varying the four factors: weighting coefficients (B_i), goodness (LG), adjustment parameter (B_o), and scoring method.

The four factors or treatments were varied in the model over different levels as listed:

coefficients - seven levels
goodness - five levels
adjustment - three levels
scoring method - two levels

The concern was to determine if any of the treatments significantly changed the mean frequency of selection of the "best" proposal.

A suitable statistical technique for determining the significance of any observed change over a number of observations of different levels of treatments is Analysis of Variance (ANOVA), (11:526).

The assumptions of ANOVA are that:

(1) The probability distributions of the dependent variables are normal.

(2) The variance of the dependent variable is constant.

(3) The samples are independent.

Regarding the assumption of normality, the variable of interest in the experiment was the number of times the "best" proposal was selected, i.e., the result of n Bernoulli trials of which the outcome was either "selected" or "not selected". The distribution of such a series of events has the binomial probability:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where n = sample size

x = the number of events of interest in n

and p = probability of occurrence of an event of interest

(11:137)

However, when the sample size n is reasonably large ($n \geq 30$), the binomial probability distribution can be approximated by a normal probability distribution (11:216), and it has been shown (14:61) that a moderate departure from normality has little effect on the test of significance of ANOVA.

A preliminary scanning of the computer model output suggested that constancy of variance was a reasonable assumption. It was decided to proceed to ANOVA on that basis and use the Cochran's "C" procedure provided with the SPSS program to test the assumption after the event (12:430).

The sample data of the computer model was statistically independent to the extent of the independence of the pseudo-random number generator. The condition of independence was regarded to be satisfied for the purposes of ANOVA for all treatments except for the treatment "method", (color or numerical scoring). For simulation of "method", the treatments were successively applied to the same sets of basic data. ANOVA was, therefore, chosen as the means by which to examine the treatment effects of "weighting coefficients", "goodness", and "adjustment parameter".

As the treatment "method" involved only two levels of treatment (color or numbers), it was appropriate to apply the t-test for population mean differences between matched samples to study the effect of "method" (11:320).

Sample Size

It was desired to have a sample size so that the ANOVA would provide information about the discrimination

of the computer model with 10 per cent confidence level with 10 per cent accuracy of estimation of the frequency of selection of the "best" proposal.

For a binomial probability distribution, the sample size can be estimated by:

$$n = \frac{Z^2(\alpha/2)}{4d^2}$$

(15:191)

where $Z(\alpha/2)$ is the two-tailed normal statistic for the desired confidence level, and

d is the difference between the true probability of selection and the estimate.

For the experimental requirements, n was calculated to be 68; 5 computer runs of 80 simulations were selected to provide an adequate data base for evaluation of results.

ANOVA Test Procedure

The computer model outputs were first tested to determine the significance of the different treatments when applied separately. SPSS procedure ONEWAY was employed. There are two steps involved in using this technique:

(1) Test the hypotheses

H_o : There is no difference in the mean proportion of proposals number 1 selected between different levels of the treatment being studied.

H_1 : There is a difference in the mean proportion of proposals number 1 selected between different levels of the treatment being studied.

The decision rule for the test is:

if $F^* \leq F(0.9; r-1, n_t-r)$, conclude H_o, otherwise conclude H_1 (11:535).

Where $F^* = \dfrac{\text{treatment mean square}}{\text{error mean square}}$

and r = the number of treatment levels

n_t = total number of observations.

The value of F* is provided as part of the SPSS output.

(2) If the test shows a difference between means, analyze the ranges within which the differences lie. Duncan's multiple range test provided in the SPSS package (12:427) is suitable for this purpose.

A multiple ANOVA analysis was then conducted of the significant treatments to examine interaction effects over the range of treatments when color scoring was used in the model.

<u>ONEWAY ANOVA Results</u>

Eight data sets were selected for ONEWAY analysis to obtain a feel for the separate treatment effects on mean

proportion of proposals number one selected. The results are presented in Table III.

The results in Table III show that, for the parameter sets tested, the treatment "weighting coefficients" was not significant at the 0.1 level in determining the frequency with which the "best" proposal was selected. Both "goodness set" and "adjustment parameter" (B_o) were significant treatments which affected the outcome of the selection.

The results of the Duncan's multiple range tests are shown in Table IV.

The results show that when the numerical scoring process was applied, the mean frequency of selection of the "best" proposal was significantly different for each goodness set of proposals. The frequency of selection of the "best" proposal increased as the quality difference between the proposals increased.

A similar result was shown when color scoring was used, except that the frequency of selection of the "best" proposal in each goodness set was consistently less than when number scoring was used and that the frequency of selection of the "best" proposal was not significantly different when the difference between the "best" and "next best" proposals was large.

TABLE III

RESULTS OF ONEWAY ANOVA TESTS

Test No.	Treatment	Levels	Fixed Parameters	Grand Mean	Cochran's C	F*	F	Result
1	coefficient set	7	B_o=0 goodness set=3 method=numbers	53.70	.627	.237	2.00	H_o
2	coefficient set	7	B_o=0 goodness set=3 method=colors	49.50	.506	1.845	2.00	H_o
3	goodness set	5	B_o=0 coeff.set=4 method=numbers	53.40	.612	50.700	2.25	H_1
4	goodness set	5	B_o=0 coeff.set=4 method=colors	50.84	.283	58.700	2.25	H_1
5	B_o	3	coeff.set=4 method=colors goodness set=2	45.30	1.000	4.026	2.81	H_1
6	B_o	3	coeff.set=4 method=colors goodness set=3	52.10	.592	15.259	2.81	H_1
7	B_o	3	coeff.set=4 method=colors goodness set=4	65.10	.705	12.010	2.81	H_1
8	B_o	3	coeff.set=4 method=colors goodness set=5	67.50	.582	3.660	2.81	H_1

TABLE IV

TABLE SHOWING HOMOGENEOUS SUBSETS OF TREATMENTS

Test No.								
1	T'ment Level	1	2	3	4	5	6	7
	Mean	52.4	53.6	52.2	53.8	54.4	54.6	55.0
	Subset	———	———	———	———	———	———	———
2	T'ment Level	1	2	3	4	5	6	7
	Mean	45.0	47.6	47.4	49.6	49.8	52.8	54.2
	Subsets	———	———	———	———	———		
			———	———	———	———	———	
					———	———	———	———
3	T'ment Level	1	2	3	4	5		
	Mean	35.0	43.8	53.8	64.8	69.8		
	Subsets	———						
			———					
				———				
					———			
						———		
4	T'ment Level	1	2	3	4	5		
	Mean	31.6	42.8	49.6	64.0	66.2		
	Subsets	———						
			———					
				———				
					———	———		
5	T'ment Level	1	2	3				
	Mean	52.0	42.8	41.0				
	Subsets	———						
			———	———				
6	T'ment Level	1	2	3				
	Mean	58.6	49.6	48.0				
	Subsets	———						
			———	———				
7	T'ment Level	1	2	3				
	Mean	69.4	64.0	62.0				
	Subsets	———						
			———	———				
8	T'ment Level	1	2	3				
	Mean	71.4	66.2	65.0				
	Subsets	———						
			———	———				

With regard to the ONEWAY analysis of the treatment "adjustment parameter" (B_o), included when color scoring was used, treatment level 1 ($B_o = +7$) was found to cause a significantly different result in the selection of the "best" proposal at all goodness levels. There was no significant difference between the effects of treatment levels 2 and 3 ($B_o = 0$ and $B_o = -7$ respectively).

The values obtained for P in the Cochran's C test show that in all cases the assumption of homogeneity of variances was met at the 0.1 level, justifying the validity of the ANOVA approach.

Difference Between Color-Scored and Numerically-Scored Results (t-Test)

The concept of value building by using numerical component scores and weights does not include the adjustment factor, B_o. It was therefore appropriate for the purpose of this test to compare numerical scores with color scores only at the $B_o = 0$ level.

The purpose of the t-test was to determine if the frequency of selection of the "best" proposal was significantly greater at the 0.1 level when number scores were used than when color scores were used.

If the mean score by numbers is M_n and the mean score by colors is M_c, then the test hypothesis is:

$$H_o : M_n = M_c$$

$$H_1 : M_n > M_c$$

and the decision procedure using SPSS output (12:271) is:

if the one-tailed probability is larger than α do not reject H_o.

The t-test was conducted over the range of goodness sets 2 to 5 and at B_i level 4 and $\alpha = .1$. The results are presented at Table V in which H_o is concluded for goodness sets 2 and 4 and H_1 for goodness sets 3 and 5.

Multiple ANOVA (MANOVA)

The ONEWAY ANOVA test results showed the effects of treatments "goodness" and "adjustment parameter" (B_o) to be significant at the 0.1 level for the fixed parameter values tested. Treatment "weighting coefficients" (B_i set) was found to be ineffective at 0.1 level of confidence for numerical scoring and $B_o = 0$ and goodness set number 3. However, when the parameter "colors" was included in the test for significance of treatment "weighting coefficients", the value of F* (1.845) was close to the value of F(2.00). It was considered advisable to include "weighting coefficients" as a treatment in the MANOVA in case it became significant at the extremes of range of treatments or when applied in conjunction with other treatments.

TABLE V

T-TEST OF RESULTS USING NUMBER AND COLOR SCORE METHODS

G'NESS SET	METHOD	MEAN	STD DEV'N	DIFFCE	T-VALUE	2-TAIL PROBY	1-TAIL PROBY	RESULT
2	Number	43.8	5.975	1.00	0.91	0.413	0.207	H_o
	Colors	42.8	6.140					
3	Number	53.8	4.382	4.20	3.02	0.039	0.020	H_1
	Colors	49.6	3.975					
4	Number	64.8	2.950	0.80	0.62	0.566	0.283	H_o
	Colors	64.0	2.915					
5	Number	70.2	5.718	4.00	4.48	0.068	0.034	H_1
	Colors	66.2	4.324					

The SPSS ANOVA sub-program is designed to handle MANOVA for factorial experimental designs.

Since the ONEWAY test for treatment "goodness set" yielded very large values of F*, it was likely that the effect of varying "goodness set" would overwhelm the effects of treatments "adjustment parameter" and "coefficient set" for goodness sets 2 through 4. A symmetrical factorial design was chosen with three levels each of:

"adjustment parameter"; $B_o = +7, 0, -7$

and "coefficient set"; set No. 1, set No. 4 and set No. 7.

The multiple classification analysis (MCA) option of the SPSS ANOVA program was used to provide an indication of the magnitude of the effect of each treatment. The outputs are presented at Appendices C1-C8.

Summary of Results of Analyses of Model Output

The ONEWAY test results show that the treatments "goodness set" and "adjustment parameter" are significant at the 0.1 level in determining the probability of selection of proposal number 1. "Coefficient set" was not a significant treatment for either number or color scoring at goodness set number 3 and adjustment parameter 2 ($B_o = 0$).

Further analysis of the treatments "coefficient set" and "adjustment parameter" taken conjointly in a

two-way MANOVA for goodness sets 2 through 4 when numerical scoring was used, show different joint effects of treatments "coefficient set" and "adjustment factor" as the level of goodness set is increased, i.e., as the difference in quality between the "best" proposal and the "second best" proposal increases.

When the difference between proposals is small, "adjustment factor" (B_o) is the significant external treatment. Large positive values of B_o increase the frequency of selection of the "best" proposal. At goodness set 2 (10% difference between proposals), B_o explained 0.25 of the selection preference, whereas the coefficient set explained only .04 of the selection preference. However, both treatments accounted for a relatively small part of the selection and a large variance of outcomes was predicted.

At goodness set 3 (20% difference between proposals), coefficient set and adjustment parameter each explained about .16 of the selection preference with still a relatively large variance of outcomes.

At goodness set 4 (30% difference between proposals), coefficient set became the dominant reason for selection preference, explaining 0.50 of the outcome while adjustment parameter explained 0.12 of the outcome.

At goodness set 5 (40% difference between proposals), coefficient set was even more dominant, explaining

0.56 while adjustment factor still explained 0.12 of the outcome. The variance of selection due to unexplained factors of the model was reduced as the gap between "best" and "next best" proposal increased.

The direction of effects of the treatments was also worthy of note. Large positive values of adjustment factor (B_o) increased the probability of selection of the "best" proposal. Negative values of B_o reduced the probability of selection. Coefficient sets with small differences in component weights forced selection toward the "best" proposal while larger differences in component weights resulted in greater variance of selection of proposals.

The results of t-tests for the effect of treatment "method" (color or numerical scoring) were less conclusive. At goodness sets 3 and 5, the discriminating power of numerical scoring, as modeled, was significantly greater than the power of color scoring at the 0.1 level. There was no significant difference between the two scoring methods at goodness sets 2 and 4.

CHAPTER IV

INTERVIEWS WITH SOURCE SELECTION PRACTITIONERS

To further examine the underlying nature of the source selection decision-making process, structured interviews were conducted with source selection practitioners and administrators in an attempt to identify their perceptions of the models in the field. The Aeronautical Systems Division's Directorate of Contracting and Manufacturing, originally established as point of contact in this research effort, provided a listing of selected ASD personnel who were at the time, or had recently been, engaged in different aspects of source selection. Thirty-one personnel were interviewed. All had been involved in at least one of the many different functions of source selection, including acquisition policy and procedure management, SSA, SSA advisor, SSAC member, SSEB Chairman, item captain, Acquisition Logistics Division (ALD) representative, program manager, principal contracting officer (PCO), and general contracting, pricing, buying and manufacturing participants.

An interview guide was prepared to ensure consistency of approach in the research. A copy of the guide is attached at Appendix D. The interviews were designed to

try to obtain an overall view of the source selection decision-making process from a participant's viewpoint and to assist in identifying the reality of the process as it is applied in practice against the theoretical models of multi-attribute decision-making. Discussions were centered around the following areas: effectiveness and efficiency of existing source selection decision procedures, relative merits of numerical and color-coding schemes of scoring the results of evaluations, influence of Contractor Inquiries (CIs) and Deficiency Reports (DRs) on the decision process, and ways of improving the source selection decision process.

Effectiveness and Efficiency of the Process

There was little agreement on what the ultimate objective of source selection should be, although the majority of the personnel interviewed agreed that existing source selection decision procedures assured the effectiveness of the process in attaining its perceived objective. Responses included the following:

> "a mechanism to appear as objective as possible in selecting a source while protecting against protests and complaints"
>
> "to get best contractor at best price"
>
> "to de-select other offers to be able to withstand protests"

"to get technically best contractor"

"to select best source for the government, all factors considered"

"to give you a good insight before you commit yourself"

"to select best supplier at best price, if price is one consideration"

"to be fair in selecting a source able to perform"

"to get the best capability in meeting the needs of the Air Force in accordance with the requirements of the solicitation".

Statements such as these clearly show a lack of agreement and possibly misunderstanding among personnel interviewed regarding the purpose/objective of the source selection process. A clear understanding of the ultimate objective of the process by its participants is essential to ensure effective evaluation of proposals and results which meet the ultimate objective of the source selection process.

Far greater agreement was found among those interviewed when asked about the effectiveness of the process in achieving the perceived overall objective. A large majority agreed that the process is usually effective in meeting its stated objectives, and that the right contractor is selected in almost all cases. Some concern was expressed though, regarding normative political override. Source selection decisions are sometimes made on political con-

siderations without adequately quantifying the risk of program failure.

A number of factors seems to hamper the effectiveness of the existing source selection process. Among factors cited was the effect that funding constraints have on the source selection decision. During the last decade, budgeting has been a major external influence on the process, creating "a temptation to make the low offer appear to meet the requirements" through extensive use of CIs and DRs.

The massive amount of data with which evaluators are confronted when evaluating proposals was seen to be a major factor in preventing a truly effective process. It was said that evaluators usually find it difficult to filter out the data in order to identify and be able to assess the key issues. It appears that source selection evaluations are being made with an excessive amount of data--far more than that which is needed--obscuring the important issues and preventing decision-makers from effectively evaluating them.

Most of the personnel interviewed expressed concern about the inefficiency of the source selection process. Many said that they considered the process to be grossly inefficient, due mainly to the large number of people involved in the evaluation stages, the excessive amount of

time taken up by evaluations, and the large amounts of data encountered in proposals. The large and detailed RFPs sent out to industry seem partly to be the cause of much of this inefficiency. The RFPs force offerors to generate large amounts of data in support of their proposals and make evaluation a time-consuming, extremely complex process which requires many evaluators in order to sort out the data.

More than half of the respondents said that the source selection process involves far too many people. Lack of expertise and evaluating experience was cited by some as contributing to the inefficiency of the process. It also appears that the government spends a disproportionately large amount of resources in obtaining a small system in relation to that which it spends in acquiring a major system. The need to streamline the process was emphasized. Some suggested that a small group of 10 to 15 qualified evaluators could reach a decision as acceptable as that made by a large number of evaluators.

Some concern was expressed regarding the amount of resources spent in areas which did not influence the final decision. Much emphasis is placed on certain areas of proposals, e.g., management. The effort evaluators put into these areas seems unwarranted when the output of such evaluations fails to have an impact on the decision

process. It was observed that there is a trend toward increasing the number of management evaluation items.

While some of the perceived inefficiency attributed to the process may be caused by the need to document everything in order to have a sound defense against potential protests, such a fear of protests appears to be unfounded. Less than 4 per cent of contracts awarded by ASD result in protests, with the majority of the protests being shown to be without foundation.

In summary, it appears that a number of factors cause many people to be involved in source selection. However, the process seems to have worked effectively, and the desired results have been achieved as well in those cases where strong management has insisted on a reduced number of evaluators.

The Scoring Process

AFR 70-15 provides broad guidance on source selection decision procedures. It discusses the use of both numerical and color-coding schemes of scoring the results of evaluations, supported by narrative statements. ASD regulations encourage the use of color-coded and narrative assessments, and numerical scoring has not been formally used in ASD since June 1972. In an attempt to identify the strengths and weaknesses of both the numerical and

color-coding techniques, personnel interviewed were asked to comment on the relative merits of each approach.

While about one-half of those interviewed expressed their preference for the use of colors, one-third indicated that both methods were equally effective in assessing proposals, with a few personnel showing a preference for the numerical scoring technique. The preference for the color-coding approach seemed to be based on the concept of providing an integrated assessment which would highlight the strengths, weaknesses, and risks of each proposal and allow the SSA greater latitude to exercise judgment.

Under the numerical scoring system, the SSA felt constrained to accept the numerical results, and a decision to select a source other than the one with the highest scoring proposal was difficult to justify. Comments were also expressed that source selection is partly a qualitative judgment process which is sometimes hard to quantify and creates difficulty in arriving at an agreed number, whereas agreement is much more easily reached using color scores. Areas such as past performance and management are sometimes difficult to weigh and score with numbers giving an unwarranted degree of precision, while color-coding provides a clearer overall picture to the decision-maker.

Individuals who expressed the view that both approaches were equally effective and would serve to accomplish essentially the same purpose, indicated that the important thing is to conduct a balanced evaluation which ensures key areas are identified appropriately and evaluated properly.

It was frequently stated that in the more objective areas, e.g., technical, evaluators made initial scores on a numerical scale. They then converted these, using cut-off values, to color scores to fit in with the source selection plan.

Those who preferred the numerical approach said that numbers provided a quicker reaction to, and identification of, slight differences between similar proposals. The numerical scoring technique appears to yield a more discrete and finer identification of differences at the attribute level; something color-coding fails to do. It forces the attribute evaluator to commit himself to a firm decision. In addition, numerical scoring allows the weighting of issues to be precisely identified in advance of scoring according to their relative importance as established in the source selection plan. Conversely, they said that color-coding introduces a degree of uncertainty and encourages political maneuvering.

In discussing numerical scoring, a variety of perceptions of the "cut-off" level of discrimination of numerical scores, one to the other, and when compared to a standard, was discovered. Some respondents said that an absolute difference between scores was a sufficient basis on which to make a decision. Most who gave an opinion said that a difference of 1 to 2 per cent between scores was significant. Less than that, other (subjective) considerations would come into the decision. About half the respondents felt they could not give an opinion, and one experienced officer said that if numerical scores were used in systems source selections, he would not consider score differences of less than 10 per cent to be significant. Respondents frequently said that in many cases ASD was concerned with buying concepts which did not lend themselves to highly objective scoring.

When evaluating proposals at the SSEB level, proposals should be compared against standards established in the solicitation document. A tendency to compare proposals with each other at this level, rather than against standards, as required by regulations (17:p.3-4) was expressed by some of those interviewed.

Contractor Inquiries and Deficiency Reports

A considerable amount of effort is spent by source

selection personnel in the preparation of CIs and DRs as the means of communicating with offerors, to provide for clarification of certain aspects of proposals, and to identify specific parts of proposals which fail to meet the government's minimum requirements. This procedure allows the offerors to correct deficiencies found by evaluators. Almost every one of the personnel interviewed agreed that although the CI/DR process is time-consuming and usually prolongs the evaluations, it is essential to obtaining a satisfactory contractual arrangement and is significant in influencing the decision process.

Responses indicated a frequent excess of CIs. This was partly due to the failure of RFPs to be definitive in some areas. The excess was also attributed to the reluctance of evaluators to make a subjective judgment, and attempts to obtain a defensible, documented position. It was suggested by some respondents that more direct talks with offerors would help to reduce the number of CIs originated and eliminate much of the paperwork created during the process. It was observed that, in those cases where ASD had used the four-step solicitation process (20:4) there was a large reduction in the use of CIs. DRs were considered to be far more critical in influencing the decision process, since these documents allow evaluators to determine how well final offers meet the govern-

ment's requirements. The most important ones are usually highlighted under "strengths and weaknesses" in evaluation reports to the higher levels of decision-making.

Although AFR 70-15 requires that proposals only be scored as originally submitted to encourage the best initial proposals, it was found that, in practice, proposals were often rescored. A review of three source selection case histories in ASD, together with responses obtained during the interviews, indicated that proposals are frequently rescored after the CI/DR process is completed. It appears that further clarification and guidance regarding rescoring of proposals may be required to ensure that a fair and consistent approach is used.

Improvements Suggested by Interviewees

It was agreed that the existing source selection process is usually effective in selecting the proposal(s) which best meets the government's cost, schedule and performance requirements, considering that what is being evaluated usually is an offeror's future performance of something which is essentially innovative. However, a great majority of the personnel interviewed saw much room for improvement of the process. The discussion that follows concentrates on those areas suggested to have the greatest potential for improvement of the overall process.

A need to integrate the source selection activity with that of preparing RFPs was expressed. A great part of the source selection plan and process is determined by the way the RFP was written. It was said that closer coordination between source selection personnel and those responsible for the preparation of RFPs would help ensure that more definitive and concise requirements go out to industry. This contact would result in more compact and precise proposals which would serve to reduce the tremendous amounts of data with which evaluators are presently being confronted, would allow the significant aspects and key issues to surface sooner, and would provide for a more effective and efficient evaluation. Some respondents suggested that the size of proposals should be controlled by defining in the RFP the number of pages of submission allowed.

Further streamlining of the process was suggested to help make it more efficient. It was suggested that a group of well-qualified and experienced personnel with broad knowledge, complemented with competent technical advisors, would result in a reduced number of evaluators and a shorter time required to assess proposals. It was said that the source selection experience of evaluators must be improved and more specific guidance provided for first-time evaluators. The lack of a viable training

program in source selection procedures for evaluators with no previous experience in source selection, makes it a difficult task for those personnel who have to learn the procedures while on the job. This shortcoming results in much unproductive time and decreased efficiency.

An awareness that many of the problems which surface during the performance of a contract are related to the contractor's data and cost tracking systems has directed an increased emphasis on the management area of proposals during evaluations. Respondents indicated that although a considerable amount of effort is spent in this area, it seldom influences the decision process. An improved approach for assessing the management area of proposals in a more realistic way was felt to be necessary, with increased emphasis being placed on a prospective contractor's past performance.

It was felt that there was a need to develop a better way of linking the cost and technical evaluations together in order to obtain a realistic cost-benefit analysis. Other suggestions in this area dealt with the need to bring together the assessments of the Cost Panel and the Contract Definitization Group at some point during the process to provide a better overall picture when considering tradeoffs between cost and technical requirements.

Increased use of the abbreviated procedures for source selection was advocated. In the abbreviated procedure a Source Selection Evaluation Committee (SSEC) assumes the responsibilities of the SSAC and SSEB. This resulted in a more efficient process. It seemed evident to some interviewees that frequently the SSAC failed to apply the judgment required in a comparative analysis of proposals, and merely served as a means of filtering the SSEB evaluation results to the SSA. It was also felt that some of the more formal requirements for source selections on lower-dollar acquisitions could be eliminated, improving the efficiency of the process without impacting on its effectiveness.

A need to rescore proposals after the DR process is completed was thought by many to be an essential procedure to ensure an optimum decision. Scoring proposals as originally submitted and as corrected seemed to be the only way to conduct a realistic appraisal.

During the course of the research, it became evident that some source selections departed significantly from the guidance provided in regulations; a fact which was felt by some people interviewed to cause some of the inefficiency attributed to the process. This was thought to be partly due to the lack of recent and current guideance in the field. AFR 70-15, the primary document for

establishing policy and procedures for the conduct of source selections in the Air Force, is now five years old, outdated and has been under revision for over a year. It was hoped that when the new issue of AFR 70-15 is published, it will provide more specific guidance for the conduct of source selections.

Some concern was expressed that major contractors have developed an ability to submit high scoring bids which makes it difficult to assess proposals which, on paper, appear to be fairly similar. Evaluation then becomes a task of determining whether the offeror is able to do what he says he can do, rather than making an objective technical decision. As this seems to be the case during many formal source selections, it becomes critical to provide the SSA with objective information on which to make a rational decision which will reduce the risk of cost overruns and program slippages.

Summary

The interviews provided a good insight of the source selection process as it is presently applied, and identified a number of difficulties perceived by source selection participants.

Even though respondents agreed that the process was effective in achieving its perceived objective, there

was little agreement as to what that objective should be. Concern was expressed regarding the inefficiency of the process. This inefficiency was attributed to the large number of people involved in source selections, the excessive amount of time taken up by evaluations, and the massive amount of data with which evaluators are confronted.

Views regarding the techniques used for scoring proposals provided a wide range of opinions of the relative merits of each approach. Preference for numerical or color scoring methods was divided.

Although it was evident that the CI and DR processes are time-consuming, they were considered to be essential and very significant in influencing the decision process and in making a satisfactory contract.

Interviewees agreed that there was room for improvement of the process. Their responses suggest some approaches for accomplishing that objective.

Source selections in ASD cover a wide range of acquisitions of varying degrees of complexity and maturity of concept. However, there is some evidence that the process is not always applied with sufficient judgment and that departures from policy and procedures occur.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

This study was directed toward identifying the process of source selection as practised in ASD. The methodology of source selection was simulated through a computer model. A perspective of the process was developed through a review of the procedural guidance and a series of interviews with ASD source selection personnel. This chapter summarizes the findings of the study and compares them with some theoretical concepts to develop a descriptive evaluation of the ASD source selection process. Finally, recommendations are made which may contribute to the improvement of the management of source selection.

Source Selection Methodology

The analysis of source selection cases in which color or narrative scoring methods were used demonstrated the possibility of evaluators incorporating an adjustment parameter (B_o) into the value building process when aggregating a group of lower-level attribute scores.

The effect of introducing negative values of B_o into the simulation model was to reduce the discrimination

of the process in selecting the "best" proposal in terms of the evaluation criteria. Positive values of B_o biased the scores in favor of the "best" proposal. The B_o effect was greatest when the difference between proposals was small. More cases of negative values of B_o were observed than positive values suggesting a "wash-out" of the component evaluations in those cases.

As might be expected from the work of Dr. Lee, the model confirmed that when the difference between the mode goodness or quality of the components of the "best" proposal and the "second best" proposal was large, the most significant internal parameter which affected the selection was the weight applied to each component (coefficient set). When the weighting difference was large, the proposal with the "best" modal quality was less likely to be selected than when weighting differences were small.

The relative effectiveness of numerical scoring and color scoring as discriminators was substantially dependent on the nature of the relative difference in the quality of proposals being compared. For some differences in quality of proposals, color-scoring provided significantly less preference for the "best" proposal than numerical scoring provided. The inconsistency of discrimination provided by color scoring is explained by the "broad banding" of the four-increment color score scale.

Evaluations of two proposals which fall on different sides of the boundary between two color bands will be discriminated by color scoring. However, if the evaluations of two proposals (which, theoretically may differ by as much as 40%) fall within the same color band, the color scoring system will not differentiate between them.

The positive features of numerical scoring when compared with color-scoring are:

(1) Absolute weights may be allocated to attributes before evaluation and scoring.

(2) The inclusion of adjustment parameters (B_o) which can wash out or bias final scores is precluded.

(3) Small differences in evaluations are recognized in the scores allocated to attributes and are discriminators of the outcome.

The disadvantages of numerical scoring are:

(1) A degree of precision of evaluation is implied which is not always realistic, particularly when dealing with the conceptual attributes of proposals.

(2) Numerical scores imply a sense of absoluteness which inhibits the exercise of qualitative judgment by the SSA.

(3) Evaluators tend to be reluctant to use the full range of scores, clustering results into a narrow band, so reducing the discriminating power of the process.

(4) It is sometimes difficult to obtain agreement on relative weights.

(5) Extreme responses are not highlighted (e.g., non-conformance).

In comparison, color scoring offers the following advantages:

(1) A convenient and powerful means by which a comparative overview of the quality of competing proposals may be visualized is provided.

(2) Subjective values of attributes may be scored with high levels of agreement.

(3) Extreme responses are highlighted.

(4) The SSA is provided with considerable scope for qualitative judgment.

The disadvantages of color-scoring are:

(1) Significant differences in objective evaluations of attributes may not be recognized in the scoring process as discriminating factors.

(2) Attributes cannot be objectively weighted to highlight comparative importance.

(3) The process permits the washing out or biasing of evaluation results by the introduction of an adjustment parameter (B_o).

Both methods of scoring have unique advantages and disadvantages. Whether one method or the other is appropriate is dependent on the nature and structure of the particular source selection involved. It is concluded from this study that the choice of the appropriate method of scoring is influenced by:

(1) The maturity of the concept being considered.

(2) The relative importance (weights) of the key attributes of the decision.

(3) The resources available to the source selection activity.

(4) The management style of the SSA.

<u>Maturity of Concept</u>

The maturity of the concept being considered strongly controls the level at which a proposal may be evaluated. When the concept is novel and the proposal is, in effect, a projection of what might be done based on broad assumptions, then the evaluation can only be realistically scored at a qualitative level. Evaluating human skills such as expectations of management or innovative capabilities is also highly conceptual and only able to

be satisfactorily expressed in qualitative terms. Conversely, when standard and predictable techniques and practices of mature concepts are being evaluated, quantitative scoring of evaluations can be done with confidence and precision. A single source selection may involve a mix of novel and mature concepts. For example, a technical area may encompass a variety of well-developed concepts, whereas the corresponding logistics area may be one in which the implications of the systemic application of the technology is entirely novel.

Weights of Attributes

In some source selections, the weight of the decision may rest heavily on a particular attribute. In others, weights of attributes may be about equal. Even at lower levels of evaluation such as the factor level, it may be necessary to weight the sub-factors to prevent the important attributes from being swamped by the many trivial attributes. Numerical scoring methods allow the use of definitive weights when needed. Color scoring is weak in its ability to reflect weightings but has the power to highlight component deficiencies when weighting is not important.

Source Selection Resources

The major resources available to a source selection activity are personnel, time and money. Personnel may be limited in numbers or specific skills. Time available may limit the depth of evaluation. Money resources may determine the extent of investigation of proposed solutions or restrict the amount of outside assistance that can be brought to bear on the source selection. All of these resource constraints may reduce both the effort that can be put into evaluating the attributes of each proposal, and the precision with which the attributes may be scored. As the potential for precision of evaluation is reduced, color scoring becomes a more suitable technique than numerical scoring.

Management Style of the SSA

Simon has written that management and decision-making may be viewed as synonymous (16:1). The management style of the SSA is an important consideration in selecting the source selection structure. The structure should provide the SSA with the kind of information he needs to be able to make an effective decision within his own frame of reference. Keen and Morton (7:62) classify decision style into five main groups:

- rational - based on analytical definition of all the variables to obtain the best decision.
- satisficing - based on effective use of available information to obtain an acceptable decision.
- procedural - based on following through standard organizational procedures toward a single outcome.
- political - based on the use of personalized bargaining between organizational units to seek an acceptable decision.
- individual - based on the decision-maker's own individual assessment of the information available to him.

This grouping of decision-making styles suggests that different decision-makers will seek different kinds of information on which to act. Rational and satisficing decision-makers are likely to feel more comfortable with numerically-scored information, whenever it may be practically applied. The procedural decision-maker is unlikely to strongly favor either numerical or color scoring, so long as he is satisfied that a correct procedure has been followed. Political and individual decision-makers are more likely to be attracted to color or narrative scoring techniques as being compatible with their own styles of management.

Choice of Scoring Method

The wide range of factors bearing on the effectiveness of a particular source selection process suggests that there is no one best technique for scoring proposals. The requirement that "a qualitative rating scale will be used in lieu of weighted scoring (1:9)" unnecessarily inhibits ASD source selection personnel from exercising the flexibility to choose the process best suited to each source selection situation.

Within a source selection case, different areas may merit different scoring processes according to the criteria discussed above. The color scoring system does not offer sufficient range to be able to satisfactorily show important differences in many areas of technical evaluation. In other areas, such as management, color scoring may be an appropriate tool when needed to indicate the outcome of largely subjective judgments. There is no overriding reason why numerical and color scoring should not be separately used in different parts of the same source selection. If done, it would present the SSA with an overview of both the objective and subjective aspects of the total evaluation. Alternatively, if it is the preference of the SSA, scores could be feasibly converted to an "all color" or "all number" presentation at the area level.

When values can be determined with higher precision than afforded by a four-increment color range, numerical scoring offers greater power of discrimination of the merits of proposals than that offered by color scoring. This advantage should not be foregone in those groups of attributes for which numerical scoring is appropriate. Other perceived problems of numerical scoring (clustering, agreement on weights, and extremes not highlighted) may be overcome by appropriate techniques. The techniques described by Beard (MAUM) and Waid (comparison matrices) provide practical and convenient ways of making objective and effective weight and score allocations with reliability and repeatability. Extreme attribute scores (mainly non-conformance) may be simply highlighted in the evaluation presentation or treated with an exclusion rule which eliminates the proposal from further analytical consideration.

Color scoring is a suitable technique for representing evaluations of subjective and highly conceptual attributes. It recognizes the imprecision inherent in such areas, yet presents a good overall comparative picture of proposals. Extremes are highlighted. Where precision of evaluation is possible, color scoring tends to wash-out significant differences. It presents problems in allocating relative weightings when weights are significant to the decision,

and allows a wide variety of outcomes because of the scope for implicit adjustment factors in the process. The disadvantages of color scoring can be minimized by management vigilance and skill in application, together with careful consideration of the suitability of areas to the application of the technique.

Procedural Aspects of Source Selection

The prime objective of the source selection process is to obtain an impartial, equitable, and comprehensive evaluation of competitive proposals which will result in the selection of a source which will offer optimum satisfaction of the government's requirements, to include cost, schedule, and performance (17:p.1-1). The wide range of conflicting responses obtained from interviewees regarding the ultimate objective of the process tends to indicate that personnel involved in source selection fail to approach the process with a common objective. This lack of agreement impacts on the quality of the final decision and reduces the overall effectiveness of the process. The need to understand and work toward a common objective in source selection cannot be overemphasized. It is essential in order to make a selection based on that objective.

AD-A105 056 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL--ETC F/G 5/10
THE SOURCE SELECTION DECISION PROCESS IN AERONAUTICAL SYSTEMS D--ETC(U)
JUN 81 C W BARCLAY, J E NIDO

UNCLASSIFIED AFIT-LSSR 12-81 NL

2 OF 2
AD
A105056

END
DATE
FILMED
11-81
DTIC

Effectiveness and Efficiency

During the evaluation of proposals, source selection personnel are confronted with vast amounts of data, a large part of which is not needed to make an effective decision in an efficient manner. This excess detracts from the decision-maker's main tasks. Providing source selection personnel with excessive amounts of data inhibits them from being able to effectively identify and assess the small amount of really important information needed to reach a decision which will result in satisfaction of the government's objectives.

The study of source selection cases during this research found examples in which areas, items or factors were broken down into many attributes for evaluation. Often there was high multi-collinearity between some attributes, indicating that they did not contribute to the decision. Some respondents to interviews expressed concern at the proliferation of sub-division of evaluation. Evaluation of management was cited as a particular area of proliferation. There appears to be a tendency, if an area is recognized as critical to the decision, to expand the sub-headings under which it is evaluated. There is the danger in this approach that proliferation of parts merely leads to an averaging of scores and obscures what

is important. The discrimination of the process is improved by keeping the number of attributes small and by applying differential weights to them according to their importance. Helman and Taylor (6:90) suggest that only three items (planning, organizing, and controlling) should be evaluated in the management area and that each item be broken into no more than four factors.

A need to develop a better way to consider cost and performance tradeoffs is suggested from interview responses. The present philosophy of source selection is to associate a cost with a technical proposal, identify acceptable proposals within a competitive cost range, and then obtain best and final offers (BAFOs). Although in theory the budget should not be a constraint, and the contract should be awarded to the offeror who best satisfies the government's requirements, in practice, budget restrictions sometimes prevent the selection of the best offeror. A tendency to emphasize cost limitation at the expense of technical feasibility may not be the best decision.

Use of Scoring Techniques

Personnel involved in source selection held a wide range of opinions on the effectiveness and appropriateness of methods of scoring proposals: numbers, colors, or

narrative. There was little objective understanding of the implications of choosing one method over another, and choice was largely a matter of subjective preference based on experience. The model experimentation conducted during this research suggests that there are circumstances in which the method of scoring proposals significantly affects the outcome.

Contractor Inquiries

There was a broad focus on the concentration of use of CIs to clarify and justify the work of evaluators beyond what was required for contract definitization. Much of the use of CIs was felt by many interviewees to be a device to protect the organization from future protests by unsuccessful offerors. Proliferation of CIs tends to extend evaluation time and can lead to technical leveling and raising low cost proposals to more favorable evaluation levels. Cases were seen where initially poorly-rated proposals were rescored to acceptable levels as the result of CI actions. It was observed that the introduction of the four-step solicitation process (20:1) contributed to a large reduction in the use of CIs.

Clearly, a balance is required between sufficient use of CIs to provide adequate contract definitization and an inefficient excess of CIs. Many personnel felt that a

proper balance was not being achieved.

<u>Problems of Source Selection</u>

The major problems confronting the conduct of source selection lie in meeting effectively and efficiently the objectives of impartiality, equitability and comprehensiveness. The descriptive model of source selection developed in this research shows it to be a highly complex process. The goals of the process are not always clearly perceived. The effect which the techniques used have on the interaction within the process are not widely understood by personnel involved in the activity. Procedural guidance tends to be fragmented and is not clear on the suitability and applicability of the techniques available to evaluators and decision authorities. The transitory nature of source selection teams precludes the development of depth of experience in many personnel key to the process.

A logical, consistent, and disciplined approach, tailored to requirements, is necessary to provide a complete and objective analysis with minimum resources. The process should efficiently communicate to the SSA a clear, complete, relevant, and objective analysis which will provide a reliable basis for the source selection decision.

Recommendations

This study points to some possible ways in which source selection may be made more effective and efficient. Many of the problems experienced in making efficient and effective source selections lie in the limited experience and understanding of the process by many of the personnel involved. Working-level expertise in source selection is limited because of the relatively short time many participants spend in the process. However, their actions impact on decisions involving very large expenditures of money and long-term operational commitments by the Air Force.

A significant contribution to improvement of the operation of the process would be to introduce short training programs for personnel entering source selections. The training should be directed toward developing:

(1) A common understanding of the Air Force objectives of source selection.

(2) A knowledge of the procedural framework of source selection.

(3) An appreciation of the scoring and weighting techniques available, their relative advantages and disadvantages, and their scope of application.

The training program should emphasize the principle of essentiality in source selection. A source selection

should be concerned with what is essential to the decision. It should focus on collecting and evaluating essential data. Efficient source selection plans should restrict the use of evaluation items and limit the factors to a few significant headings which will facilitate meaningful discrimination between proposals. The literature review, findings, and discussions in this study provide some insights of the source selection process which provide a basis upon which a suitable training program could be built.

In parallel with training development, a review of source selection procedures is advisable. The following changes are recommended:

(1) Avoid directed use of specific scoring or weighting techniques.

(2) Encourage source selection planning tailored to the specific characteristics of the acquisition.

(3) Facilitate integration of RFP development with source selection planning.

(4) Provide guidance on relating cost evaluation to technical evaluation.

The better understanding of procedures and policy, and the objectivity that can flow from such measures should result in smaller, more purposeful source selection teams,

and more powerful decision support mechanisms for the SSA.

The model developed here is an attempt to describe the complexities of the source selection decision process in ASD. It is not complete, and does not purport to be so. However, it is hoped that it will provide a useful basis from which to improve understanding of the process. There is scope for much further work, some of which is suggested in the preceding pages.

APPENDICES

APPENDIX A

LISTING OF COMPUTER PROGRAM

```
      PROGRAM SOUSIM
      SIMULATION OF A SOURCE SELECTION DECISION CEL. BY BARCLAY/NICO
      INTEGER KY(1:5),LG,JV,ID,IS,N,SEED,J,K,L,C
      REAL GD(1:5,1:5),B(1:7,1:3),RO(1:3),Y(1:5),AX,TE,FX,K(1:5,1:5,1:5,
     1    1:30),A,PERC(1:5,1:7,1:3),XA,KB
      CHARACTER MARK*80,CODE*7
      DATA MARK/'
     1                                          '/
      GOODNESS MATRIX GD(LG,JV) : LG=GOODNESS SET, JV=SID IN SET
      DO 88 LG=1,5
          GD(LG,1)=10.
 88   CONTINUE
          GD(1,2)=10.
      GD(2,2)=9.
      GD(3,2)=8.
      GD(4,2)=7.
          GD(5,2)=6.
          GD(1,3)=9.
      GD(2,3)=8.
      GD(3,3)=7.
      GD(4,3)=5.
      DATA GD(5,3),GD(1,4),GD(2,4),GD(3,4),GD(4,4),GD(5,4),GD(1,5),GD(2,
     15)/
     15.,3.,7.,3.,5.,4.,7.,6./
      GD(3,5)=5.
      GD(4,5)=4.
      GD(5,5)=3.
      MATRIXES GD AND COEFFICIENTS B:ID=80 SET,IS=3 SET,N=3 COEFFICIENT
      B(1,1)=.5
      DO 488 N=2,5
          B(1,N)=.1
 488  CONTINUE
      DATA B(2,1),B(2,2)/.5,.2/
      DO 588 N=3,5
          B(2,N)=.1
 588  CONTINUE
      DATA B(3,1),B(3,2)/.4,.3/
      DO 688 N=3,5
          B(3,N)=.1
 688  CONTINUE
      DATA B(4,1),B(4,2),B(4,3),B(4,4),B(4,5),B(5,1),B(5,2),B(5,3),B(5,4
     1),B(5,5)/
     1.4,.2,.2,.1,.1,.3,.3,.2,.1,.1/
      B(6,1)=.3
      DO 788 N=2,4
          B(6,N)=.2
 788  CONTINUE
      B(6,5)=.1
      DO 888 N=1,5
          B(7,N)=.2
 888  CONTINUE
      DATA RO(1),RO(2),RO(3)/7.,1.,-7/
          DO 60 C=1,2
              SEED=91257
              IF(C.EQ.1) THEN
                  CODE='COLORS'
                  ELSE
```

```
                    CODE='NUMBERS'
                ENDIF
                    DO 80 L=1,5
C GENERATE RANDOM-BASED VALUES OF X FOR EACH GOODNESS LEVEL
      DO 999 LG=1,5
          DO 1000 JV=1,5
              DO 1100 N=1,5
                  DO 1200 I=1,80
      M=GO(LG,JV)
      SEED=MOD((2-298*SEED+9991),199017)
          FX=SEED/199017.
      XA=SQRT(FX 10*M)
          XO=(2+SQRT(4-.4*(10**FX-M*FX+M)))/.2
          XF=(2-SQRT(4-.4*(10**FX-M*FX+M)))/.2
          IF(XO.GT.M.AND.XO.LE.10.)THEN
          XB=XO
          ENDIF
          IF(XF.GT.M.AND.XF.LE.10.)THEN
          XB=XF
          ENDIF
          IF(XA.LE.M)THEN
          AX=XA
          ELSE
            AX=XB
          ENDIF
                  IF(C.EQ.1)THEN
          IF(AX.LT.1.25)THEN
          X(LG,JV,N,I)=0.
          ELSEIF (AX.LT.4.25)THEN
          X(LG,JV,N,I)=2.5
          ELSEIF (AX.LT.8.)THEN
          X(LG,JV,N,I)=6.
          ELSE
              X(LG,JV,N,I)=10.
          ENDIF
                      ELSE
                      X(LG,JV,N,I)=AX
                      ENDIF
1200              CONTINUE
1100          CONTINUE
1000      CONTINUE
999   CONTINUE
      DO 1399 IO=1,3
          DO 1488 IS=1,7
              DO 1588 LG=1,5
                  DO 1688 JV=1,5
                  KY(JV)=0
1600              CONTINUE
      DO 1788 I=1,80
      TE=0
          DO 1589 JV=1,5
      Y(JV)=GO(IC)+(1-GO(IC)/10)*(B(IS,1)*X(LG,JV,1,I)+B(IS,2)*X(LG,JV,2
     1,I)+B(IS,3)*X(LG,JV,3,I)+B(IS,4)*X(LG,JV,4,I)+B(IS,5)*X(LG,JV,5,I)
     1)
          IF(Y(JV)-TE.GT.0.02-TE)THEN
          TE=Y(JV)
          ENDIF
```

```
1300        CONTINUE
            DO 1900 JV=1,5
            IF(Y(JV).EQ.TE)THEN
            KY(JV)=KY(JV)+1
            ENDIF
1900        CONTINUE
1700  CONTINUE
            PERC(LG,IS,ID)=KY(1)*100./(KY(1)+KY(2)+KY(3)+KY(4)+KY(5))
      IF(C.EQ.1.AND.L.EQ.1)THEN
      PRINT*
      PRINT '(/T10,A,T45,F4.1)','RESULT OF 80 SIMULATIONS OF BO=',BO(ID)
      PRINT '(T10,A,T40,F5.2,T45,F5.2,T50,F5.2,T55,F5.2,T60,F5.2)',
     *'AND COEFFICIENT SET',B(IS,1),B(IS,2),B(IS,3),B(IS,4),B(IS,5)
      PRINT '(T10,A,T45,I3)','AND GOODNESS SET OF BIDS',IS
      PRINT '(/T5,A,T30,A)','HIGHEST VALUE','NO. OF OCCURRENCES'
            DO 2000 JV=1,5
            PRINT '(T12,A,T13,I2,T35,I3)','Y',JV,KY(JV)
2000        CONTINUE
      PRINT '(//T10,A)','HISTOGRAM OF FREQUENCY OF SELECTION OF BIDS'
      PRINT*
            DO 2100 JV=1,5
            PRINT '(T10,A)','I'
            PRINT '(T5,A,T6,I2,T10,A,T11,A)','Y',JV,'I',MARK(1:KY(JV))
2100        CONTINUE
      PRINT '(T10,A)','I'
      PRINT*
      PRINT '(T10,A,T47,F7.2)','PER CENT OF TIME BID NO.1 SELECTED=',PER
     *C(LG,IS,ID)
      PRINT '(T15,A,T36,F5.2)','LAST HIGHEST Y WAS',TE
      ELSE
            GOTO 1500
      ENDIF
1500              CONTINUE
1400        CONTINUE
1300  CONTINUE
3000  PRINT '(/////T22,A)','RESULT OF 80 SIMULATIONS OF BID DATA'
      PRINT '(/T17,A)','PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE'
      PRINT '(/T27,A,T38,I2,T44,A)','RUN NUMBER',L,CODE
      PRINT '(///T5,A,T20,A,T34,A)','COEFFICIENT','BO','GOODNESS SET'
      PRINT '(T9,A)','SET'
      PRINT '(T17,A,T25,A,T31,A,T42,A,T53,A,T64,A,T75,A)',
     *'I','I','1','2','3','4','5'
      PRINT '(T17,A,T25,A)','I','I'
      PRINT '(T3,77A)',('*',K=1,77)
      DO 3000 IS=1,7
            DO 3100 ID=1,3
            PRINT '(T17,A,T25,A)','I','I'
            PRINT '(T10,I2,T17,A,T19,F4.1,T25,A,T28,F6.1,
     *      T39,F6.1,T50,F6.1,T61,F6.1,T72,F6.1)',
     *      IS,'I',BO(ID),'I',PERC(1,IS,ID),PERC(2,IS,ID),
     *      PERC(3,IS,ID),PERC(4,IS,ID),PERC(5,IS,ID)
3100        CONTINUE
            PRINT '(T17,A,T25,A)','I','I'
      PRINT '(T3,77A)',('*',K=1,77)
3000  CONTINUE
 0                CONTINUE
 0              CONTINUE

      END
```

APPENDIX B

SUMMARY TABLES OF COMPUTER OUTPUT

RESULT OF 90 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 1 COLORS

COEFFICIENT SET	SG	GOODNESS SET 1	2	3	4	5
1	7.0	34.5	52.7	37.1	66.3	55.5
1	0.0	23.7	46.9	33.3	60.5	63.9
1	-7.0	27.5	46.7	31.4	59.8	63.9
2	7.0	34.8	57.5	48.9	72.0	72.0
2	0.0	30.7	50.0	35.7	61.0	64.5
2	-7.0	28.4	48.8	34.5	61.0	63.4
3	7.0	37.8	53.8	54.9	71.6	71.1
3	0.0	30.2	48.5	45.9	66.7	72.0
3	-7.0	30.2	47.7	44.0	64.5	72.0
4	7.0	39.6	59.1	54.9	71.5	74.4
4	0.0	32.2	50.0	43.4	64.2	70.4
4	-7.0	30.2	46.9	45.9	63.0	70.4
5	7.0	39.6	58.8	59.9	70.6	73.5
5	0.0	31.5	47.1	53.1	63.5	69.9
5	-7.0	31.5	45.8	51.2	64.3	69.9
6	7.0	43.8	54.0	59.7	77.5	75.2
6	0.0	36.5	45.2	50.4	61.0	72.4
6	-7.0	36.5	44.0	49.4	59.8	72.9
7	7.0	37.0	45.7	52.1	62.9	75.3
7	0.0	34.9	47.7	51.1	60.7	75.4
7	-7.0	35.5	43.2	47.3	63.5	75.9

RESULT OF 30 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 2 COLORS

COEFFICIENT SET	EC	GOODNESS SET 1	2	3	4	5
1	7.0	35.0	52.9	54.7	55.6	63.1
1	0.0	31.0	43.5	43.4	51.1	51.7
1	-7.0	30.6	43.0	43.4	51.1	61.0
2	7.0	35.3	54.9	55.5	61.2	72.3
2	0.0	30.2	46.1	45.1	56.5	65.7
2	-7.0	30.2	46.1	45.9	54.1	66.7
3	7.0	35.6	59.0	54.3	64.4	72.3
3	0.0	31.3	51.2	42.9	52.3	63.1
3	-7.0	29.8	51.2	43.7	49.4	63.1
4	7.0	35.2	59.0	53.5	66.7	72.3
4	0.0	33.7	46.2	44.4	62.5	63.1
4	-7.0	32.5	47.0	44.4	61.3	63.1
5	7.0	41.4	54.2	55.4	70.7	75.3
5	0.0	29.8	50.0	47.0	60.5	63.1
5	-7.0	29.6	50.0	45.4	59.3	63.1
6	7.0	42.7	50.6	55.1	66.5	81.7
6	0.0	33.7	41.4	50.5	62.7	72.9
6	-7.0	32.6	41.4	50.5	62.7	71.4
7	7.0	35.1	41.1	51.7	67.9	77.4
7	0.0	34.7	43.5	52.9	67.0	74.7
7	-7.0	31.3	43.5	52.9	65.9	73.3

RESULT OF 30 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 3 COLORS

COEFFICIENT SET	LO	GOODNESS SET 1	2	3	4	5
1	7.0	35.0	53.6	55.3	62.7	53.4
1	0.0	31.9	39.0	50.0	59.0	55.1
1	-7.0	28.4	39.0	49.2	58.5	55.3
2	7.0	39.3	50.6	57.5	69.5	61.4
2	0.0	28.9	44.6	50.4	62.2	55.5
2	-7.0	28.9	43.4	48.7	59.8	55.8
3	7.0	35.4	51.2	53.1	66.3	71.5
3	0.0	24.1	44.0	50.9	65.9	55.0
3	-7.0	24.1	43.4	45.9	64.6	55.4
4	7.0	35.2	43.0	55.6	69.9	57.9
4	0.0	29.1	40.7	51.7	66.3	60.5
4	-7.0	26.7	39.5	43.9	65.9	59.3
5	7.0	40.0	53.5	55.9	70.4	70.4
5	0.0	35.4	45.3	51.9	65.4	53.4
5	-7.0	25.3	44.7	51.8	64.2	52.2
6	7.0	39.5	53.6	57.5	73.5	75.0
6	0.0	29.3	37.8	55.6	72.3	55.3
6	-7.0	25.2	36.3	53.7	71.5	55.3
7	7.0	31.5	47.4	54.7	70.0	53.1
7	0.0	31.5	47.7	50.7	70.9	55.7
7	-7.0	25.5	42.5	57.3	69.0	55.5

RESULT OF 99 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 4 COLORS

COEFFICIENT	CC	GOODNESS OF FIT 1	2	3	4	5
1	7.0	35.3	47.7	57.5	54.2	61.4
1	0.0	27.1	42.0	51.9	48.9	57.3
1	-7.0	25.7	41.5	51.9	48.8	55.1
2	7.0	35.7	46.3	57.5	61.0	65.3
2	0.0	24.1	40.2	55.7	50.9	58.5
2	-7.0	24.1	39.0	52.9	49.4	55.5
3	7.0	35.1	45.5	57.5	67.5	73.2
3	0.0	25.2	34.6	52.4	63.9	59.3
3	-7.0	25.2	34.6	52.4	62.5	59.3
4	7.0	40.2	50.0	61.9	67.9	71.1
4	0.0	29.0	40.2	54.9	60.0	62.2
4	-7.0	25.9	38.3	52.4	58.8	61.0
5	7.0	38.6	43.0	55.5	69.5	72.3
5	0.0	29.8	35.3	51.9	62.2	64.5
5	-7.0	29.9	35.3	51.9	61.0	64.5
6	7.0	33.1	40.9	52.7	71.5	74.7
6	0.0	31.8	37.3	54.9	66.3	70.4
6	-7.0	31.8	34.6	54.9	63.8	69.1
7	7.0	33.0	38.5	53.4	69.4	73.0
7	0.0	35.5	37.1	51.4	67.1	75.5
7	-7.0	33.3	37.1	53.7	63.4	72.0

RESULT OF 50 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 5 COLORS

COEFFICIENT SET	BC	GOODNESS SET 1	2	3	4	5
1	7.0	42.2	39.6	55.7	54.3	57.5
1	0.0	35.6	36.5	46.7	60.2	53.0
1	-7.0	35.6	36.8	47.1	59.0	55.5
2	7.0	40.5	42.5	59.8	71.1	72.3
2	0.0	33.7	35.6	47.7	61.4	53.3
2	-7.0	33.7	35.5	47.4	61.0	53.3
3	7.0	44.6	46.1	57.4	73.5	53.7
3	0.0	31.3	36.0	47.1	65.9	65.1
3	-7.0	31.3	36.0	47.5	64.6	62.2
4	7.0	45.1	50.0	61.7	65.2	71.1
4	0.0	35.4	35.2	43.9	64.7	63.7
4	-7.0	34.1	32.2	49.4	61.4	65.3
5	7.0	44.7	46.7	62.7	70.2	75.7
5	0.0	33.3	38.3	44.5	62.7	64.5
5	-7.0	32.1	35.5	44.5	61.4	64.5
6	7.0	43.2	53.4	59.3	70.5	81.5
6	0.0	39.3	43.2	50.9	62.7	70.0
6	-7.0	37.4	41.4	43.1	62.7	57.5
7	7.0	44.3	41.1	45.7	65.5	73.5
7	0.0	42.2	40.9	47.7	69.4	73.2
7	-7.0	35.0	36.9	47.7	69.4	70.7

RESULT OF 80 SIMULATIONS OF BID DATA

PER-CENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 1 NUMBERS

COEFFICIENT SET	CO	GOODNESS SET 1	2	3	4	5
1	7.0	43.8	57.5	56.3	66.3	73.0
1	0.0	27.5	50.0	47.5	56.3	63.8
1	-7.0	26.3	45.0	43.8	56.3	65.0
2	7.0	45.3	62.5	51.3	75.0	72.5
2	0.0	27.5	52.5	53.8	61.3	63.8
2	-7.0	25.0	48.8	43.8	58.8	67.5
3	7.0	47.5	60.0	58.8	78.8	75.0
3	0.0	32.5	45.0	53.8	65.0	63.8
3	-7.0	30.0	45.0	55.0	62.5	65.0
4	7.0	50.0	65.0	67.5	78.8	77.5
4	0.0	35.3	48.8	53.8	65.0	63.8
4	-7.0	35.0	43.8	52.5	65.0	67.5
5	7.0	51.3	63.8	70.0	81.3	82.5
5	0.0	35.3	47.5	57.5	66.3	72.5
5	-7.0	35.0	46.3	58.8	62.5	63.8
6	7.0	52.5	62.5	70.0	80.0	83.8
6	0.0	41.3	47.5	55.0	67.5	75.0
6	-7.0	38.8	41.3	53.8	67.5	72.5
7	7.0	52.5	63.8	63.8	83.8	81.3
7	0.0	36.3	53.8	53.8	66.3	77.5
7	-7.0	33.8	50.0	53.8	65.0	72.5

RESULT OF 80 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 2 NUMBERS

COEFFICIENT SET	TU	GOODNESS SET 1	2	3	4	5
1	7.0	40.0	60.0	57.5	67.5	73.8
1	0.0	30.0	51.3	51.3	58.8	75.0
1	-7.0	23.8	48.8	50.0	57.5	72.5
2	7.0	41.3	61.3	60.0	72.5	78.8
2	0.0	26.3	53.8	48.8	62.5	77.5
2	-7.0	23.8	53.8	45.0	62.5	75.0
3	7.0	35.0	63.8	50.0	76.3	83.8
3	0.0	23.8	47.5	40.0	62.5	75.0
3	-7.0	23.8	45.0	40.0	60.0	75.0
4	7.0	41.3	52.5	50.0	76.3	81.3
4	0.0	30.0	43.8	45.0	67.5	77.5
4	-7.0	27.5	46.3	45.0	62.5	75.0
5	7.0	43.8	65.0	57.5	77.5	85.0
5	0.0	27.5	42.5	45.0	66.3	73.8
5	-7.0	26.3	38.8	42.5	65.0	75.0
6	7.0	41.3	60.0	51.3	61.3	85.0
6	0.0	36.3	47.5	43.8	70.0	80.0
6	-7.0	35.0	42.5	40.0	66.3	77.5
7	7.0	45.0	51.3	53.8	62.5	85.0
7	0.0	32.5	38.8	47.5	[illegible]	73.8
7	-7.0	30.0	32.5	45.0	68.8	77.5

RESULT OF 80 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 3 NUMBERS

COEFFICIENT SET	BC	GOODNESS SET 1	2	3	4	5
1	7.0	73.8	57.5	60.0	65.0	67.5
1	0.0	35.3	53.4	50.0	53.9	53.8
1	-7.0	35.0	51.3	53.0	50.0	50.0
2	7.0	42.5	60.0	65.0	72.5	63.8
2	0.0	35.3	47.5	52.5	65.0	60.0
2	-7.0	35.0	47.5	43.8	60.0	58.8
3	7.0	47.5	61.3	65.0	76.3	72.5
3	0.0	35.0	47.5	56.3	67.5	67.5
3	-7.0	32.5	47.5	55.0	63.8	63.8
4	7.0	45.0	62.5	56.3	77.5	72.5
4	0.0	33.5	46.3	53.3	65.0	55.0
4	-7.0	30.0	46.3	55.0	62.5	63.8
5	7.0	48.8	62.5	55.0	77.5	77.5
5	0.0	32.5	48.8	57.5	66.8	55.3
5	-7.0	32.5	47.5	57.5	68.8	65.0
6	7.0	52.5	65.0	71.3	62.5	55.1
6	0.0	31.3	50.0	57.5	70.0	71.0
6	-7.0	31.3	47.5	56.3	66.8	65.0
7	7.0	51.3	66.3	75.3	53.8	63.3
7	0.0	32.4	55.0	51.3	73.4	57.5
7	-7.0	33.4	52.5	57.5	73.8	57.5

RESULT OF 80 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 4 NUMBERS

COEFFICIENT SET	RO	GOODNESS SET 1	2	3	4	5
1	7.0	40.0	45.0	50.0	63.8	66.3
1	0.0	28.8	41.3	57.5	56.3	60.0
1	-7.0	26.3	38.8	55.0	56.3	58.8
2	7.0	42.5	43.8	63.8	66.3	68.8
2	0.0	28.8	33.8	56.3	56.3	63.8
2	-7.0	27.5	32.5	55.0	56.3	63.8
3	7.0	43.8	46.3	62.5	72.5	75.0
3	0.0	30.0	37.5	55.0	58.8	65.0
3	-7.0	28.8	35.0	55.0	56.3	65.0
4	7.0	43.8	53.8	66.3	71.3	73.8
4	0.0	35.0	38.8	55.0	60.0	65.0
4	-7.0	33.8	33.8	53.8	60.0	65.0
5	7.0	43.8	43.8	67.5	72.5	76.3
5	0.0	35.0	37.5	57.5	66.3	72.5
5	-7.0	31.3	36.3	56.3	65.0	71.3
6	7.0	47.5	50.0	65.0	73.8	73.8
6	0.0	36.3	38.8	56.3	70.0	73.8
6	-7.0	35.0	38.8	55.0	67.5	68.8
7	7.0	51.3	53.8	73.8	80.0	81.3
7	0.0	33.8	36.3	57.5	71.3	72.5
7	-7.0	33.8	35.0	55.0	71.3	72.5

RESULT OF 30 SIMULATIONS OF BID DATA

PERCENT OF TIME BID NUMBER 1 HAD HIGH VALUE

RUN NUMBER 5 NUMBERS

COEFFICIENT	[illegible]	GOODNESS SET 1	2	3	4	5
1	7.0	48.8	52.5	57.5	66.3	75.3
1	0.0	35.0	42.5	55.0	63.3	57.5
1	-7.0	33.8	41.3	51.3	62.5	53.8
2	7.0	53.8	55.0	67.5	70.0	77.5
2	0.0	36.3	33.8	56.7	65.0	58.8
2	-7.0	32.5	36.3	51.3	65.0	57.5
3	7.0	48.8	52.5	66.7	71.3	73.8
3	0.0	32.5	40.0	53.8	62.5	71.3
3	-7.0	31.3	37.5	50.0	60.0	63.3
4	7.0	55.0	61.3	63.8	73.8	81.3
4	0.0	40.0	38.3	55.8	66.3	73.8
4	-7.0	35.0	35.0	52.5	66.3	72.5
5	7.0	51.3	57.5	62.5	76.3	81.3
5	0.0	35.0	35.0	52.5	66.3	71.0
5	-7.0	33.8	35.0	48.8	66.3	63.8
6	7.0	32.5	56.3	63.8	77.5	82.5
6	0.0	40.0	50.0	55.0	73.8	80.0
6	-7.0	37.5	43.8	52.5	71.3	73.8
7	7.0	57.5	58.8	66.0	80.0	82.5
7	0.0	37.5	45.3	61.3	70.0	72.5
7	-7.0	35.0	43.8	46.7	67.5	72.5

APPENDIX C

OUTPUTS OF MANOVA TESTS

APPENDIX C1

GOODNESS SET 2

* * * * * * * * A N A L Y S I S   O F   V A R I A N C E * * * * * * * *

CHOICE
BY BI
BO

* *

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	438.533	4	109.633	4.047	.008
BI	62.400	2	31.200	1.152	.327
BO	376.133	2	188.067	5.943	.006
2-WAY INTERACTIONS	171.467	4	42.867	1.582	.200
BI BO	171.467	4	42.867	1.582	.200
EXPLAINED	610.000	8	76.250	2.815	.016
RESIDUAL	975.200	36	27.089		
TOTAL	1585.200	44	36.027		

45 CASES WERE PROCESSED.
0 CASES (0 PCT) WERE MISSING.

APPENDIX C2

GOODNESS SET 3

* * * * * * * * A N A L Y S I S   O F   V A R I A N C E * * * * * * * * *

CHOICE
BY BI
BO

* *

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	666.667	4	166.667	4.544	.006
BI	329.733	2	164.867	4.495	.018
BO	336.933	2	168.467	4.593	.017
2-WAY INTERACTIONS	199.733	4	49.933	1.361	.267
BI BO	199.733	4	49.933	1.361	.267
EXPLAINED	866.400	8	108.300	2.953	.012
RESIDUAL	1320.400	36	36.678		
TOTAL	2186.800	44	49.700		

45 CASES WERE PROCESSED.
0 CASES (0 PCT) WERE MISSING.

APPENDIX C3

GOODNESS SET 4

```
* * * * * * * * *  A N A L Y S I S   O F   V A R I A N C E  * * * * * * * *
                CHOICE
             BY BI
                BO
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	919.422	4	229.856	15.889	.001
BI	747.378	2	373.689	25.831	.001
BO	172.044	2	86.022	5.945	.006
2-WAY INTERACTIONS	61.422	4	15.356	1.061	.390
BI BO	61.422	4	15.356	1.061	.390
EXPLAINED	980.844	8	122.605	8.475	.001
RESIDUAL	520.800	36	14.467		
TOTAL	1501.644	44	34.128		

45 CASES WERE PROCESSED.
0 CASES (0 PCT) WERE MISSING.

APPENDIX C4

GOODNESS SET 5

* * * * * * * * A N A L Y S I S   O F   V A R I A N C E * * * * * * * *
CHOICE
BY BI
BO
* *

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	1305.422	4	326.355	23.109	.001
BI	1081.111	2	540.555	38.277	.001
BO	224.311	2	112.156	7.942	.001
2-WAY INTERACTIONS	41.822	4	10.456	.748	.571
BI BO	41.822	4	10.455	.743	.571
EXPLAINED	1347.244	8	168.405	11.925	.001
RESIDUAL	508.400	36	14.122		
TOTAL	1855.644	44	42.174		

45 CASES WERE PROCESSED.
0 CASES (PCT) WERE MISSING.

APPENDIX C5

GOODNESS SET 2

* * M U L T I P L E   C L A S S I F I C A T I O N   A N A L Y S I S * *

CHOICE
BY BI
BO

GRAND MEAN = 44.13

VARIABLE + CATEGORY	N	UNADJUSTED DEV'N	UNADJUSTED ETA	ADJUSTED FOR INDEPENDENTS DEV'N	ADJUSTED FOR INDEPENDENTS BETA	ADJUSTED FOR INDEPENDENTS + COVARIATES DEV'N	ADJUSTED FOR INDEPENDENTS + COVARIATES BETA
BI							
1	15	.40		.40			
2	15	1.20		1.20			
3	15	-1.60		-1.60			
			.20		.20		
BO							
1	15	4.00		4.00			
2	15	-1.27		-1.27			
3	15	-2.73		-2.73			
			.49		.49		
MULTIPLE R SQUARED					.277		
MULTIPLE R					.526		

APPENDIX C6

GOODNESS SET 3

* * M U L T I P L E   C L A S S I F I C A T I O N   A N A L Y S I S * *

CHOICE
BY BI
BO

* *

GRAND MEAN = 50.93

VARIABLE + CATEGORY	N	UNADJUSTED DEV'N	UNADJUSTED ETA	ADJUSTED FOR INDEPENDENTS DEV'N	ADJUSTED FOR INDEPENDENTS BETA	ADJUSTED FOR INDEPENDENTS + COVARIATES DEV'N	ADJUSTED FOR INDEPENDENTS + COVARIATES BETA
BI							
1	15	-3.73		-3.73			
2	15	1.13		1.13			
3	15	2.60		2.60			
			.39		.39		
BO							
1	15	3.80		3.80			
2	15	-1.27		-1.27			
3	15	-2.53		-2.53			
			.39		.39		
MULTIPLE R SQUARED					.305		
MULTIPLE R					.552		

APPENDIX C7

GOODNESS SET 4

* * M U L T I P L E   C L A S S I F I C A T I O N   A N A L Y S I S * *

CHOICE
BY BI
BO

* *

GRAND MEAN = 53.69

VARIABLE + CATEGORY	N	UNADJUSTED DEV'N	UNADJUSTED ETA	ADJUSTED FOR INDEPENDENTS DEV'N	ADJUSTED FOR INDEPENDENTS BETA	ADJUSTED FOR INDEPENDENTS + COVARIATES DEV'N	ADJUSTED FOR INDEPENDENTS + COVARIATES BETA
BI							
1	15	-5.69		-5.69			
2	15	2.04		2.04			
3	15	3.64		3.64			
			.71		.71		
BO							
1	15	2.64		2.64			
2	15	-.62		-.62			
3	15	-2.02		-2.02			
			.34		.34		
MULTIPLE R SQUARED					.612		
MULTIPLE R					.782		

APPENDIX C8

GOODNESS SET 5

* * M U L T I P L E   C L A S S I F I C A T I O N   A N A L Y S I S * *

CHOICE
BY BI
BO

* *

GRAND MEAN = 67.31

VARIABLE + CATEGORY	N	UNADJUSTED DEV'N	UNADJUSTED ETA	ADJUSTED FOR INDEPENDENTS DEV'N	ADJUSTED FOR INDEPENDENTS BETA	ADJUSTED FOR INDEPENDENTS + COVARIATES DEV'N	ADJUSTED FOR INDEPENDENTS + COVARIATES BETA
BI							
1	15	-6.11		-6.11			
2	15	.22		.22			
3	15	5.89		5.89			
			.75		.75		
BO							
1	15	3.09		3.09			
2	15	-.98		-.98			
3	15	-2.11		-2.11			
			.35		.35		
MULTIPLE R SQUARED					.703		
MULTIPLE R					.839		

APPENDIX D

GUIDE TO INTERVIEWS WITH SOURCE SELECTION PRACTITIONERS

GUIDE TO INTERVIEWS WITH SOURCE SELECTION PRACTITIONERS

Introduction

Outline the topic and scope of the study to the interviewee.

Obtain details of the interviewee's background and experience in source selection.

Specific Points of the Discussion

The purpose of this section of the interview is to obtain the interviewee's perceptions of:

(1) The ultimate objective of the source selection process and the effectiveness and efficiency of the process toward achieving the objective.

(2) The comparison between numerical and color scoring methods and comments on their relative merits. The significant level of discrimination in numerical scoring systems.

(3) The significance of Contractor Inquiries (CIs) and Deficiency Reports (DRs) to the decision process. Is the effort commensurate with the usefulness of CIs and DRs?

(4) Changes that could be made to improve the source selection process.

Closing Discussion

Invite additional comment on the source selection process which might aid the research.

SELECTED BIBLIOGRAPHY

A. REFERENCES CITED

1. Aeronautical Systems Division, Air Force Systems Command. Source Selection Guide. ASDP 800-7. Wright-Patterson AFB OH, 1978.

2. Beard, Major Robert J., USAF. "The Application of Multi-Attribute Utility Measurement (MAUM) to the Weapon Systems Source Selection Process." Unpublished research report No. 0140-80, Air Command and Staff College, Maxwell AFB AL, 1980.

3. Dawes, R.M. "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making," American Psychologist, 1971, pp. 180-88.

4. DeWispelare, Aaron, A.P. Sage, and C.C. White, III. "A Multicriterion Planning Aid for Defense Systems Acquisition with Application to Electronic Warfare Retrofit," Proceedings of the Ninth Annual DOD/FAI Acquisition Research Symposium, United States Naval Academy, Annapolis MD. June 1980.

5. Dycus, Bob. "Improving the Source Selection Process by Measuring the Human Response of Proposal Evaluators." Proceedings of the Sixth Annual Department of Defense Procurement Research Symposium, Army Procurement Research Office, U.S. Army Logistics Management Center, Fort Lee VA. June 1977.

6. Helman, Theodore, LtCol, USAF and Robert L. Taylor, Maj, USAF. "A Conceptual Model for Evaluating Contractor Management During Source Selection." National Contract Management Journal, Vol 10, Number 2, Winter 1976-77, pp.88-108.

7. Keen, Peter G.W. and Micheal S. Scott Morton. Decision Support Systems - An Organizational Perspective. Reading: Addison-Wesley, 1978.

8. Lee, David A. "Sensitivity of Offerors' Scores to Variations in Item Weights and Item Scores." Proceedings of the Seventh Annual Acquisition Research Symposium, Hershey PA. June 1978.

9. Logistics Management Institute. Briefings on Defense Procurement Policy and Weapon Systems Acquisition. Washington, 1969.

10. Milligan, Captain John N., USAF. "A Critical Appraisal of Source Selection Procedures." Unpublished master's thesis. AFIT/GSM/SM/79S-10, AFIT/EN, Wright-Patterson AFB OH, September 1979. AD A076158.

11. Neter, John, William Wasserman, and G.A. Whitmore. Applied Statistics. Boston: Allyn and Bacon, 1978.

12. Nie, Norman H., and others. Statistical Package for the Social Sciences. 2d ed. New York: McGraw-Hill, 1975.

13. Peterson, Steven W. "Numerical Methods for the Evaluation of Potential Research and Development Contractors." Unpublished research report. USAMC-ITC-02-08-75-214, USAMC Internal Training Center, Red River Army Depot, Texarkana TX, April 1975.

14. Scheffe, Henry. The Analysis of Variance. New York: Wiley, 1959.

15. Shannon, Robert E. Systems Simulation - The Art and Science. Englewood Cliffs: Prentice-Hall, 1975.

16. Simon, Herbert A. Administrative Behaviour. New York: Macmillan, 1957.

17. U.S. Department of the Air Force. Source Selection Policy and Procedures. AFR 70-15. Washington: Government Printing Office. 16 April 1976.

18. U.S. Department of Defense. Defense Acquisition Regulations. Washington: Government Printing Office.

19. U.S. Department of Defense. Major System Acquisition Procedures. DOD Directive 5000.2. Washington: Government Printing Office, March, 1980.

20. U.S. Department of Defense. Selection of Contractual Sources for Major Defense Systems. DODD 4105.62. Washington: Government Printing Office, January, 1976.

21. Waid, Charles C. "Determining Value: A Process for Quantifying Subjective Beliefs." Proceedings of the Seventh Annual Acquisition Research Symposium, Hershey PA. June 1978.

22. Williams, Robert F. "Problems in Numerical Input for the Source Selection Decision." Defense Systems Management Review, Vol 3, Number 3, Summer 1980, pp.122-128.

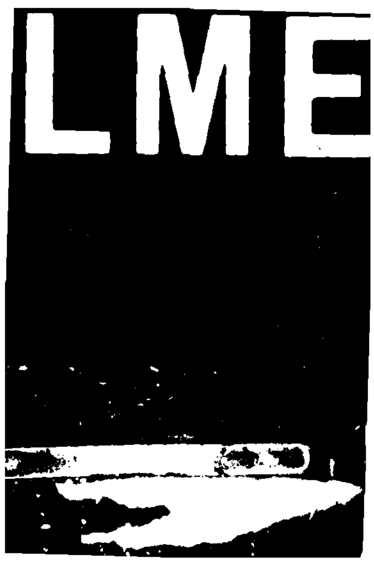
B. RELATED SOURCES

Aeronautical Systems Division, Air Force Systems Command. _The Source Selection Process_. Wright-Patterson AFB OH, 15 Jan 1978.

Proceedings of the Eighth Annual DOD/FAI Acquisition Research Symposium. Naval War College, Newport RI, May 1979.

Shaw, Major Graham, USAF. "Source Selection Process Handbook for the Air Force Space Division." Unpublished research report No. 2185-80, Air Command and Staff College, Maxwell AFB AL, 1980.

U.S. Department of Defense. _Major System Acquisitions_. DOD Directive 5000.1. Washington: Government Printing Office, March 1980.


LME